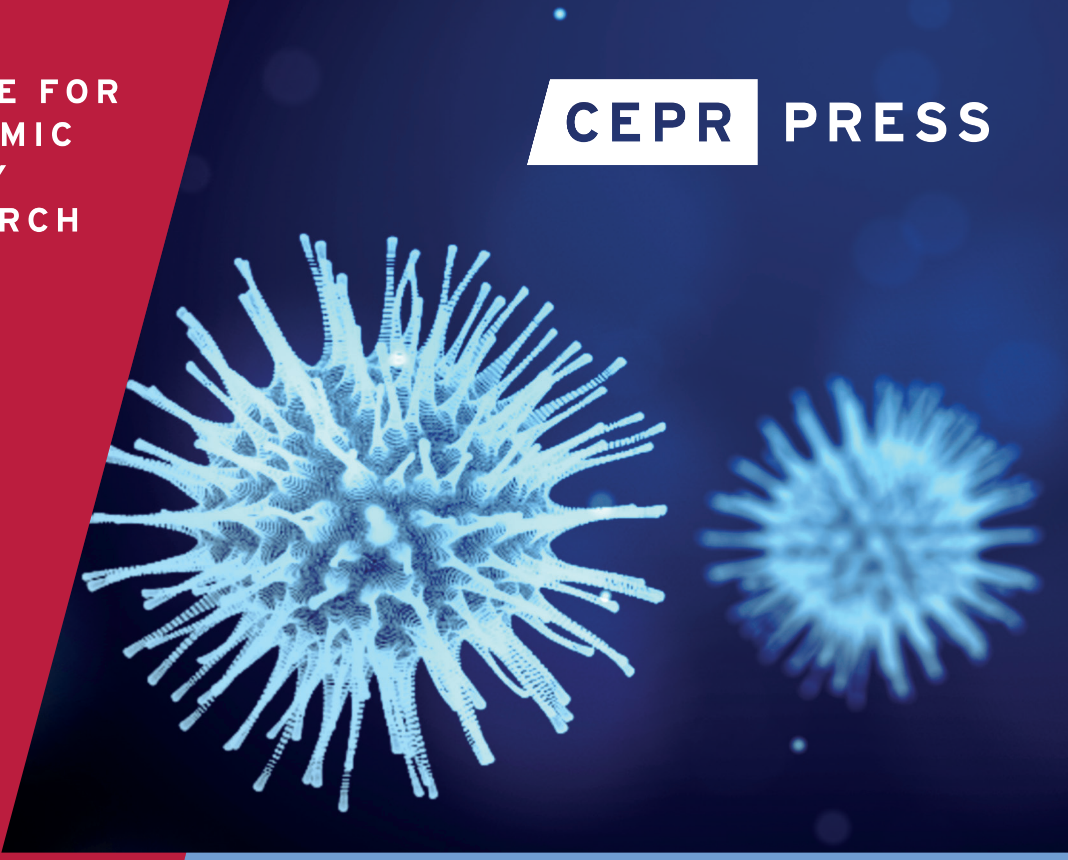
CENTRE FOR
ECONOMIC
POLICY
RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 65
20 JANUARY 2021

VACCINE POLICIES
Nikhil Vellodi and Joshua Weiss

SCHOOL OPENINGS
Emanuele Amodio, Michele Battisti, Andros Kourtellos, Giuseppe Maggio and Carmelo Massimo Maida

COSTS AND BENEFITS OF LOCKDOWNS
Anna Scherbina

SHORT SELLING
Levy Schattmann, Jan-Oliver Strych and Joakim Westerholm

INVESTMENT FUNDS
Daniel Carvalho and Martin Schmitz

VACCINATION INTENT
Jacques Bughin, Michele Cincera, Kelly Peters, Dorota Reykowska, Marcin Zyszkiewicz and Rafal Ohmed

Covid Economics
Vetted and Real-Time Papers

Covid Economics, Vetted and Real-Time Papers, from CEPR, brings together formal investigations on the economic issues emanating from the Covid outbreak, based on explicit theory and/or empirical evidence, to improve the knowledge base.

Founder: Beatrice Weder di Mauro, President of CEPR
Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics. Other queries should be sent to covidecon@cepr.org.

The Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,500 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society. CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Journal, Applied Economics
American Economic Journal, Economic Policy
American Economic Journal, Macroeconomics
American Economic Journal, Microeconomics
American Economic Review
American Economic Review, Insights
American Journal of Health Economics
Canadian Journal of Economics
*Econometrica**
Economic Journal
Economics of Disasters and Climate Change
International Economic Review
Journal of Development Economics
*Journal of Econometrics**
Journal of Economic Growth
Journal of Economic Theory
*Journal of the European Economic Association**
Journal of Finance
Journal of Financial Economics
Journal of Health Economics
Journal of International Economics
*Journal of Labor Economics**
Journal of Monetary Economics
Journal of Public Economics
Journal of Public Finance and Public Choice
Journal of Political Economy
Journal of Population Economics
Quarterly Journal of Economics
*Review of Corporate Finance Studies**
Review of Economics and Statistics
*Review of Economic Studies**
Review of Financial Studies

(*) Must be a significantly revised and extended version of the paper featured in *Covid Economics*.

Covid Economics
Vetted and Real-Time Papers

Issue 65, 20 January 2021

Contents

Optimal vaccine policies: Spillovers and incentives[1]

Nikhil Vellodi[2] and Joshua Weiss[3]

Date submitted: 7 January 2021; Date accepted: 17 January 2021

We offer a novel theoretical framework to study optimal vaccination policies. The key features of the model are that agents: 1) differ both in their potential exposure (x) to others and vulnerability (y) to severe illness, 2) exert negative externalities through interaction, and 3) can take voluntary preventative measures, for instance self-isolation. Our main result is a complete characterization of the second-best policy. Three striking features emerge. First, it is non-monotone – people with intermediate y are vaccinated more than those with either low or high y. Second, it exhibits an exposure premium among those who do not self-isolate – people with higher x require lower overall risk, xy, to be vaccinated. Third, for those who voluntarily self-isolate, it is invariant to y, depending only upon x. Numerical results demonstrate that policies vaccinating only the most vulnerable perform significantly worse than other simple heuristics, especially when supplies are limited.

1 We thank Miquel Oliu-Barton, Francis Bloch, Paula Onuchic, Bary Pradelwski, Ludovic Renou, Evan Sadler and Olivier Tercieux for helpful comments.

2 Paris School of Economics.

3 Federal Reserve Bank of Richmond and the Institute for International Economic Studies (IIES), Stockholm University.

CEPR PRESS

Covid Economics 65, 20 January 2021: 1-46

1 Introduction

Vaccine policies are a crucial ingredient in the eradication of pandemics. Given the scarcity of supply, and the urgent need for roll-out, the prudent prioritization of vaccine allocation is of paramount importance. Prevailing views revolve around two key considerations, vulnerability risk and transmission risk. Sensible arguments can be made for allocating the vaccine to those most at risk from suffering severe adverse symptoms conditional upon infection, as well as those who would likely spread the disease through a high rate of contact with others. (see Section 5 for a detailed discussion of these varying approaches) Thus far, however, little attention has been devoted to the role that incentives and behavior play in designing optimal vaccine policy. A recent report by the World Health Organization documents the growing importance and policy relevance of various behavioral incentives couched under the term "lockdown fatigue", referring to the rejection of mandatory isolation requirements.[1] As such, prudent policy design should account for agency in the design of second-best policies, rather than mandating rules and assuming they are fully implemented.

A burgeoning literature aims to incorporate behavior into standard *susceptible-infected-recovered* (SIR) dynamic epidemiological models, primarily to study the evolution of contagion rates.[2] Such models offer a strong methodology for generating robust quantitative predictions, but often become analytically unwieldy when combined with other considerations such as multi-dimensional heterogeneous agents. This complexity is necessary to ascertain the relative importance of vulnerability versus transmission risk in vaccine prioritization policies.

To this end, we offer a simple, tractable model that abstracts from dynamic considerations found in SIR models, but admits multi-dimensional heterogeneity. The model comprises a unit measure of individuals, indexed by both their potential exposure to other agents, x, and their vulnerability, y. Formally, x denotes an individual's probability of entering an *interaction pool*, wherein infection occurs through pairwise interactions, and y denotes their likelihood of suffering

[1]See https://tinyurl.com/y43jpxne.

[2]See Funk et al. (2010) for a comprehensive review.

Covid Economics 65, 20 January 2021: 1-46

a severe reaction to the disease *conditional* upon contracting it. As a consequence, holding y fixed, an individual with a higher x not only faces a higher risk of an adverse outcome, but also exerts a greater negative externality on others by increasing the overall size of the interaction pool. Finally, upon receiving an interaction opportunity, an individual may incur a fixed cost c to self-isolate instead.[3]

Our main result – Propositions 1 and 2 combined – provides a full, analytic characterization of the optimal allocation of vaccines taking as given incentives to voluntarily self-isolate, i.e. the *second-best* policy. It displays three striking features. First, it is *non-monotone* – people with intermediate y are vaccinated more than those with either low or high y. Such individuals over-interact relative to the social optimum, so they are particularly costly to society when unvaccinated. Second, among those not self-isolating, the optimal policy exhibits an *exposure premium* – people with higher x require a lower overall risk threshold xy to be vaccinated. This feature reflects the social value of vaccinating individuals that tend to infect others. Third, for those who self-isolate, the policy depends only upon x and not upon y. The latter two properties reveal two different ways that x is more relevant for vaccine allocation than y. For those who self-isolate, y loses its significance relative to x because individuals face no risk of infection yet their cost of avoiding interaction is increasing in the likelihood of interaction, x. For those who do not self-isolate, x and y both contribute equally to their overall risk, but only x contributes to the negative externality they impose on others. Grouping agents by behavior, the second and third features describe *within-group* allocation, while the first feature describes *between-group* allocation. On this final point, since the interacting group contains those behaving inefficiently, it is this group that receives an increased quantity of vaccines, thus generating the non-monotonicity. Beyond this main contribution, we provide further analytical insights. For instance, by characterizing the first-best allocation (vaccine allocation combined with a mandatory lockdown rule) – Proposition 3 – we highlight the key distortions that private incentives entail, most importantly socially excessive interaction by agents with intermediate y.

[3]We extend the analysis to other forms of behavior, such as vaccine hesitancy.

Covid Economics 65, 20 January 2021: 1-46

In addition to these theoretical findings, in Section 5 we perform various numerical exercises, not only to further explore key properties of the optimal policy, but also to compare it to commonly considered heuristics, such as: 1) y-policies that vaccinate only the most vulnerable, 2) x-policies that vaccinate only the most interactive, 3) xy-policies that vaccinate according to effective risk, and 4) random policies that do not correlate vaccines with any observable characteristics. We consider various performance criteria, including (utilitarian) welfare, death rates, isolation costs, and a novel measure that computes the additional quantity of vaccine required to render a policy welfare-equivalent to the second-best. This final measure informs the decision between taking time to evaluate and implement optimal policies and implementing off-the-shelf heuristics.[4]

Several policy-relevant insights emerge. Most importantly, y-policies perform significantly worse than other heuristics, and this performance gap is largest when vaccine supply is heavily constrained or the disease is particularly contagious or deadly. In those cases, y-policies end up targeting only self-isolating individuals and generate no spillovers. This result is of significant practical relevance, as many countries adopt such policies at the start of vaccine roll-out, precisely when supply is at its most limited. Beyond this, we find that vaccine allocation should be geographically dispersed rather than concentrated and should favor regions that either do not or struggle to implement mandatory lockdown policies. Finally, as vulnerability and exposure become more negatively correlated – an empirically relevant case – the welfare losses from heuristic policies relative to the second-best are exacerbated. These results can only obtain in settings that explicitly account for both heterogeneous vulnerability and exposure, as we do here. As such, we believe our results, both analytical and numerical, will help provide qualitative insights into hitherto under-explored features of vaccine allocation.

1.1 Related Literature

Our paper lies at the intersection of two separate approaches to modeling policy interventions during epidemics: the *networks*-based approach and the *dynamic contagion*-based approach. In

[4]Such a trade-off is relevant in the case of COVID-19 since many vaccines remain unused as of the writing of this paper. Moreover, most of the currently available vaccines expire relatively rapidly.

the networks-based approach, there is a literature in computer science that studies the problem of targeted interventions on random / adaptive networks.[5] These papers typically abstract from distortions to optimal policies due to equilibrium behavior. Within economics, papers such as Galeotti et al. (2020) and Ballester et al. (2006) explicitly consider incentives and behavior in network models. In particular, Galeotti et al. (2020) introduces a far richer network of interactions, and a far more general payoff structure, with leading applications to supply chains and social media, where local effects are strong. In contrast, our model involves random matching, and thus cannot capture such local effects. That said, random matching is an established approach to modeling epidemics – it forms the core of SIR-type models – where interactions with passing acquaintances are an important form of transmission. Finally, these papers do not have multi-dimensional types, while the tractability of our model allows us to obtain stronger characterizations. On this note, our model of interaction combines random matching with heterogeneous and partially endogenous contact rates, but crucially also spillover effects – an individual's contact rate is determined by the overall market thickness, itself determined by other individuals' contact rates. To the best of our knowledge, it constitutes in itself a theoretical contribution to the search and matching literature. In particular, our paper is related to Farboodi et al. (2020), who introduce heterogeneous contact rates into a decentralized asset market with random matching and also study how agency costs shape the endogenous distribution of contact rates.

The dynamic contagion-based approach typically takes the well-established SIR model of dynamic contagion and enriches it with agent incentives and behavior.[6] These papers are primarily concerned with making quantitative, dynamic predictions as well as evaluating inter-temporal policy trade-offs. Such concerns are crucial ingredients when making quantitative forecasts. We argue that agent heterogeneity is equally important, and thus propose a simple framework to explore this complementary theme while abstracting from complex dynamics.

Our work relates to the "risk compensation" phenomenon, which describes how mitigating the

[5] See Pastor-Satorras and Vespignani (2002), Shaw and Schwartz (2008), Gross et al. (2006), Epstein et al. (2008).

[6] See Brotherhood et al. (2020), Greenwood et al. (2019), and Geoffard and Philipson (1996).

Covid Economics 65, 20 January 2021: 1-46

downsides to risky actions might not reduce the prevalence of adverse outcomes if it is outweighed by greater risk-taking.[7] Such papers typically study how responses to policy alter the effectiveness of interventions, typically finding that they might mitigate efficacy. In contrast, we focus on how incentives can re-direct the targeting of interventions between the different characteristics of agents, a feature we believe is novel to this literature.

Finally, of most pressing applied concern is recent work considering the design of vaccine policies, and in particular in response to the COVID-19 epidemic. A large pre-COVID-19 literature exists looking at vaccination policy and equilibrium behavior, as summarized in Chen and Toxvaerd (2014). Our paper shares certain similarities with these, in particular demonstrating how second-best outcomes exhibit inefficiently low self-isolation or vaccine uptake. To the best of our knowledge, ours is the first paper to explicitly consider both margins of heterogeneity, and their interaction with optimal vaccination policies. Regarding COVID-19 in particular, Babus et al. (2020) consider a framework with multi-dimensional agents, but absent both contact externalities and behavioral considerations. Pathak et al. (2020) take a very different approach, bringing ethical considerations into a multi-priority approach. Finally, Rowthorn and Toxvaerd (2020) insert vaccination policy into a dynamic SIR framework.

2 Baseline Model

We first introduce the baseline model without behavioral considerations. A unit measure of agents are indexed by two-dimensional type $(x, y) \in [0, 1]^2$ drawn from a continuously differentiable distribution F with full support and with density function f. Each agent (x, y) is placed in an *interaction pool* with probability x. Once in the pool, they meet another agent with probability μ, where μ is the mass of agents in the pool. If an interaction occurs, each agent contracts the virus with probability $\alpha > 0$, and conditional upon infection, an agent (x, y) dies with probability y. Thus, we refer to x as an agent's *exposure type* and y as an agent's *vulnerability type*. Agents not placed in the pool cannot contract the virus. Each agent receives a terminal payoff of 0 if

[7] See Talamás and Vohra (2020), Kremer (1996), and Greenwood et al. (2019) for an extensive discussion.

Covid Economics 65, 20 January 2021: 1-46

they survive, and $-b < 0$ if they die. Denote by $\lambda \triangleq \alpha\mu$ the *transmission rate*, so that $x\lambda$ is the likelihood that type (x, y) contracts the virus. Simple algebra confirms that an agent's ex-ante expected payoff is

$$u(x, y) = -b\lambda xy. \tag{1}$$

Note that $u(x, y)$ is decreasing in λ, x, y and crucially is determined by the product xy – which we refer to as the agent's *risk-type* – rather than each component separately.

Vaccination Policy – A social planner aims to maximize welfare by allocating a fixed supply $\beta \in (0, 1)$ of vaccine.[8] If an agent is vaccinated, any interaction involving that agent cannot cause infection, i.e. $\alpha = 0$ for this agent's interactions. Formally, a *vaccine policy* is a mapping $v : [0, 1]^2 \to [0, 1]$ where $v(x, y)$ is the probability that an individual with type (x, y) is vaccinated and such that v is feasible:

$$\int_0^1 \int_0^1 v(x, y) f(x, y) dx dy \leqslant \beta. \tag{2}$$

Under policy v, the transmission rate becomes

$$\lambda = \alpha \int_0^1 \int_0^1 \big(1 - v(x, y)\big) x f(x, y) dx dy \tag{3}$$

and ex-ante individual expected utility is

$$u(x, y) = -b\lambda xy\big(1 - v(x, y)\big).$$

Since the distribution over types, F, has a pdf, we can restrict attention to vaccination policies such that for all (x, y), $v(x, y) \in \{0, 1\}$.

Voluntary Self-Isolation – We now extend the model by considering behavioral incentives and their implications for optimal policy. We first focus on incentives to voluntarily self-isolate and later turn to incentives to accept vaccination given possible adverse side-effects. Upon receiving an opportunity to enter the interaction pool, an agent can now pay a cost $c > 0$ to avoid doing so.

[8] In this baseline, this is identical to minimizing total deaths.

Covid Economics 65, 20 January 2021: 1-46

A *strategy profile* is a mapping $\sigma : [0,1]^2 \to [0,1]$, where $\sigma(x,y)$ is the probability an agent with type (x,y) does not pay the cost. The transmission rate is now a function of agents' strategies:

$$\lambda = \alpha \int_0^1 \int_0^1 \sigma(x,y)\big(1 - v(x,y)\big)xf(x,y)dxdy, \tag{4}$$

where we leave the dependence of strategies and the transmission rate on the vaccination policy implicit. Note that an individual's interaction strategy can but need not depend explicitly on whether they are vaccinated; all types are vaccinated either with probability 0 or probability 1, so conditioning on an individual's type is sufficient.

Equilibrium – Upon receiving an opportunity to enter the interaction pool, an unvaccinated individual with type (x,y) faces the optimization problem

$$\max\{-c, -\lambda yb\},$$

where the first term is the cost of isolating and the second term is the expected cost of death that results from interacting. An unvaccinated individual's optimal strategy is thus

$$\sigma^*(x,y) \begin{cases} = 1 & y < y^* \\ \in [0,1] & y = y^* \\ = 0 & y > y^*, \end{cases}$$

where

$$y^* \triangleq c/(\lambda b). \tag{5}$$

In words, each unvaccinated individual self-isolates and avoids the interaction pool if they are sufficiently vulnerable. All individuals use the same threshold y^* and their strategies do not depend on the frequency with which they receive interaction opportunities, x. Given σ^*, an

Covid Economics 65, 20 January 2021: 1-46

unvaccinated individual gets utility

$$u(x,y) = \begin{cases} -b\lambda x y & y < y^* \\ -cx & y \geq y^*. \end{cases}$$

A vaccinated individual has no reason to avoid interaction, always interacts when given the opportunity, and gets utility 0. Equilibrium is defined in the usual fashion for anonymous large games: given the aggregate distribution over actions, each agent is best-responding, while the distribution itself is consistent with the individual strategies of agents.

Definition 1. Fix $v \in V$. A *v-equilibrium* is a collection $\{\sigma, \lambda\}$ such that equations 4 and 5 are satisfied.

Since externalities in our model are purely negative, the induced game is one of strategic substitutes and consequently admits a unique equilibrium.

Lemma 1. *For all $v \in V$, there exists a unique v-equilibrium.*

Given a policy v, define welfare $\mathcal{W}$ to be the integral over all agents' utility:

$$\mathcal{W} = \int_0^1 \int_0^1 u(x,y) f(x,y) dx dy, \tag{6}$$

where we leave the dependence of welfare on vaccination policy implicit.

2.1 Model Discussion

The separation of an agent's unconditional risk type into their likelihood of interaction, x, and their likelihood of death conditional on infection, y, is crucial to better understand their distinct roles. The marginal distribution $F(\cdot, y)$ models a network of interactions in a tractable manner, in the spirit of random graphs but taken in the mean-field limit.[9] The assumption that F is smooth is

[9] Specifically, $F(\cdot, y)$ resembles the *degree distribution* of a random graph, but is technically different insofar as matching is modelled as random and bilateral in our setting.

Covid Economics 65, 20 January 2021: 1-46

made for simplicity, and is not necessary for the qualitative insights we provide. The model could easily be generalized to allow that the contact rate within the interaction pool be $\Theta(\mu)$, where Θ is a continuous, strictly increasing function. The idea that viral contraction occurs simply through pairwise interaction is a reduced form for an un-modelled presence of contagious agents already present within the interaction pool. It is also designed to resemble standard SIR models, in an attempt to bridge the two approaches. Equation 1 captures cleanly the fundamental notion that an agent's overall risk of death is the product of their rate of exposure, their conditional likelihood of death, and aggregate behavior.

Regarding self-isolation, the idea that an agent pays the cost to avoid interaction only after receiving an interaction opportunity gives rise to y-ordered preferences over self-isolation, as x linearly scales both the benefit of avoiding interaction and the cost. This gives rise to the natural prediction that more vulnerable people are more likely to self-isolate, regardless of their interaction rates. It also implies that the planner prefers to vaccinate high x individuals because any strategy they play while unvaccinated is more costly. We can alternatively model self-isolation by only allowing individuals to pay the interaction avoidance, c, prior to receiving any interaction opportunities. In this case, the cost of self-isolation does not scale with x and the planner is more inclined to vaccinate individuals with high xy rather than just high x.

3 Optimal Policy

3.1 Policy Examples

Before proceeding to solve for second- and first-best policies, it is instructive to define some natural policies, some of which form the backbone of real-world policies. In Section 5, we compare their performance to optimal policies. Denote by V the set of all policies.

Definition 2. A policy $v \in V$ is *monotone on* $S \subset [0,1]^2$ if $v(x,y) = 1$ and $(x,y) \in S$ implies that $v(x',y') = 1$ for all $(x',y') \in S$ such that $x' \geqslant x$ and $y' \geqslant y$. A policy $v \in V$ is *monotone* if $v(x,y) = 1$ implies that $v(x',y') = 1$ for all (x',y') such that $x' \geqslant x$ and $y' \geqslant y$.

Covid Economics 65, 20 January 2021: 1-46

CEPR PRESS

We now describe some leading examples of monotone policies.

Definition 3. A monotone policy v is a *y-policy* if $v(x, y) = 1$ for some x implies that $v(x', y') = 1$ for all $y' \geqslant y$, $x' \in [0, 1]$. A monotone policy v is an *x-policy* if $v(x, y) = 1$ for some y implies that $v(x', y') = 1$ for all $x' \geqslant x$, $y' \in [0, 1]$.

These policies give priority to a particular type dimension. Arguably the most common policy sought in practice is a y-policy, which would allocate the vaccine only to the most vulnerable, while x-policies aim more toward achieving *herd immunity*.

Definition 4. A monotone policy v is *risk ranking* or an *xy-policy* if $v(x, y) = 1$ implies $v(x', y') = 1$ for all (x', y') such that $x'y' \geqslant xy$.

Risk ranking policies are characterized by an *iso-risk* threshold such that agents are vaccinated if and only if their risk type, xy, is greater than the threshold.

Definition 5. A policy $v \in V$ is an *x-threshold policy* if there exists an *x-threshold function* $h : [0, 1] \to [0, 1]$ such that $v(x, y) = 1$ if and only if $x \geqslant h(y)$. A policy $v \in V$ is a *y-threshold policy* if there exists a *y-threshold function* $h : [0, 1] \to [0, 1]$ such that $v(x, y) = 1$ if and only if $y \geqslant h(x)$.

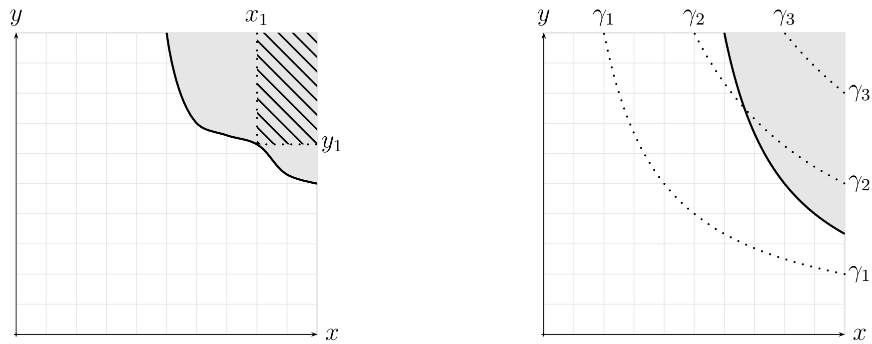
Note that every monotone policy is both an x- and y-threshold policy with a decreasing threshold function. For example, an xy-policy is characterized by the x-threshold function $h(y) = A/y$; risk, xy, is constant along the boundary. In the rest of the paper, we often refer to x- or y-threshold functions as just threshold functions when the difference is clear given the context.

Definition 6. A policy $v \in V$ that is monotone on S exhibits an *exposure premium on S* if, on S, it is characterized by a y-threshold function $h(\cdot)$ such that $xh(x)$ is strictly decreasing in x. A monotone policy $v \in V$ exhibits an *exposure premium* if it has a y-threshold function $h(\cdot)$ such that $xh(x)$ is strictly decreasing in x.[10]

[10] An equivalent definition exists in terms of x-threshold functions.

Covid Economics 65, 20 January 2021: 1-46

Figure 1: Policy Examples



Left panel: monotone policy. Black line: threshold function. Dashed area: $\{(x, y)|x \geqslant x_1, y \geqslant y_1\}$. Grey shaded area: $v(x, y) = 1$. Right panel: exposure premium policy. Black line: x- or y-threshold function. Dotted lines: xy-policies with iso-risk thresholds $\gamma_1 = 0.2, \gamma_2 = 0.5, \gamma_3 = 0.8$. Grey shaded area: $v(x, y) = 1$.

In words, if a monotone policy exhibits an exposure premium, then individuals with higher x are vaccinated with lower levels of risk overall, xy. Thus, the policy places a premium on an individual's exposure, x, relative to their vulnerability, y.

To illustrate the definition of exposure premium, consider the vaccination boundary under a monotone policy traced out by a differentiable y-threshold function $h(\cdot)$. Along the boundary, the marginal rate of substitution of x for y (MRS) is $-h'(x)$. On the other hand, the MRS along an iso-risk curve is y/x. It is readily checked that the exposure premium property is equivalent in this case to the MRS being greater along the policy boundary than along an iso-risk curve.[11] Thus, a policy exhibits an exposure premium if it is more willing to substitute x for y than simply ranking by risk would imply. We might expect this to hold when a policy places a premium on vaccinating individuals that tend to infect others.

[11]To wit, take $(x_0, h(x_0))$ on the boundary. Then the MRS of the iso-risk curve that intersects h at x_0 is $h(x_0)/x_0$. The exposure premium property thus boils down to checking that $h(x_0) + x_0 h'(x_0) < 0$.

3.2 Second-Best Policy

We now turn to the main question of the paper: what is the optimal vaccine policy taking incentives into account. That is, what is the *second-best policy*? Given the equilibrium strategies uncovered in equation 5, the second-best problem reduces to choosing $y^* \in [0,1]$, $\lambda \in [0,1]$, and $v : [0,1]^2 \to [0,1]$ to maximize the Lagrangian

$$\begin{aligned}\mathcal{L} =& \mathcal{W} + \gamma_1 \left(\beta - \int_0^1 \int_0^1 v(x,y) f(x,y) dx dy\right) + \gamma_2 \left(y^* - \min\left\{\frac{c}{\lambda b}, 1\right\}\right) \\ &+ \gamma_3 \left(\lambda - \alpha \int_0^{y^*} \int_0^1 (1 - v(x,y)) x f(x,y) dx dy\right),\end{aligned} \tag{7}$$

where $\mathcal{W}$ is given by equation (6) and the constraints are the vaccine supply constraint, the incentive constraint derived from equation (5), and the transmission rate constraint derived from equation (4).

We begin by showing that the planner's optimal vaccination policy is an x-threshold policy with a threshold function that is decreasing in y below y^* and constant above y^*.

Proposition 1. *The planner's problem is equivalent to one in which they choose a decreasing continuous function $g : [0,1] \to [0,1]$ and real numbers $g^* \in [0,1]$, $y^* \in [0,1]$, and $\lambda \in [0,1]$ to maximize the Lagrangian*

$$\begin{aligned}\mathcal{L} =& \mathcal{W} + \gamma_1 \left(\beta - \int_0^{y^*} \int_{g(y)}^1 f(x,y) dx dy - \int_{y^*}^1 \int_{g^*}^1 f(x,y) dx dy\right) + \gamma_2 \left(y^* - min\left\{\frac{c}{\lambda b}, 1\right\}\right) \\ &+ \gamma_3 \left(\lambda - \alpha \int_0^{y^*} \int_0^{g(y)} x f(x,y) dx dy\right),\end{aligned}$$

where

$$\mathcal{W} = -\lambda b \int_0^{y^*} \int_0^{g(y)} y x f(x,y) dx dy - c \int_{y^*}^1 \int_0^{g^*} x f(x,y) dx dy.$$

To understand this result, note that the interaction threshold, y^*, is the individually optimal threshold for voluntary self-isolation. As such, the second-best policy discriminates between agents depending on their incentives to self-isolate. Above y^*, individuals self-isolate and the value of

Covid Economics 65, 20 January 2021: 1-46

vaccination is in avoiding the cost associated with self-isolation. This cost scales with x, but is invariant to y. Thus, the optimal policy is an x-policy. Below y^*, individuals interact and face the risk of infection, so the cost of leaving them unvaccinated scales with y. Moreover, the private cost of infection and the public spillover cost through λ both scale with x. Thus, the optimal policy is a monotone policy with a decreasing x-threshold function $g(\cdot)$; a higher y implies a higher benefit of vaccination, so a lower x is allowed.

Note that the constraints $\lambda \in [0,1]$ and $y^* \in [0,1]$ do not bind as long as the constraints multiplying the Lagrangian multipliers γ_2 and γ_3 hold. To understand the planner's problem in more detail, it is useful to consider the first derivatives of the Lagrangian.

First, for any $y \leq y^*$,

$$\frac{\partial \mathcal{L}}{\partial g(y)} = \left[-\lambda b y g(y) + \gamma_1 - \gamma_3 \alpha g(y)\right] f\left(g(y), y\right). \tag{8}$$

The marginal benefit of increasing $g(y)$ is the value of relaxing the vaccine supply constraint, γ_1. The marginal cost is the sum of two components. An individual with type $(g(y), y)$ – who interacts even when unvaccinated since $y \leq y^*$ – dies with probability $\lambda y g(y)$, which generates utility loss b. Second, the measure of unvaccinated interacting individuals increases by $g(y)$, which tightens the equilibrium transmission rate constraint by $\alpha g(y)$ at cost γ_3.

Next, for the threshold g^* that prevails for $y > y^*$, we have

$$\frac{\partial \mathcal{L}}{\partial g^*} = (\gamma_1 - c g^*) \int_{y^*}^{1} f(g^*, y) dy. \tag{9}$$

As before, the marginal benefit of increasing the threshold g^* is the value of relaxing the vaccine supply constraint, γ_1. In this case, the marginal cost is the increase in interaction avoidance costs, cg^*, since individuals with type $y > y^*$ prefer to self-isolate when unvaccinated.

Covid Economics 65, 20 January 2021: 1-46

Covid Economics 65, 20 January 2021: 1-46

For the transmission rate λ, we have

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \begin{cases} -b \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy + \gamma_2 \frac{c}{\lambda^2 b} + \gamma_3 & \lambda > c/b \\ -b \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy + \gamma_3 & \lambda < c/b \end{cases} \tag{10}$$

The marginal benefit of increasing λ is the value of relaxing the equilibrium transmission rate constraint, γ_3, as well as the value of relaxing the incentive compatibility constraint, γ_2, multiplied by the effect of a change in λ on the individually optimal threshold y^*. The marginal cost is an increase in the transmission and therefore death rate for interacting individuals. When $\lambda < c/b$, the individually optimal threshold y^* is equal to 1, so small changes in λ have no effect on y^*.

Finally, for the interaction threshold y^*, if $y^* < 1$, we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial y^*} =& -\lambda b y^* \int_0^{g(y^*)} xf(x,y^*)dx + c\int_0^{g^*} xf(x,y^*)dx \\ & - \gamma_1 \left(\int_{g(y^*)}^1 f(x,y^*)dx - \int_{g^*}^1 f(x,y^*)dx \right) + \gamma_2 - \gamma_3 \alpha \int_0^{g(y^*)} xf(x,y^*)dx \\ =& c\int_{g(y^*)}^{g^*} xf(x,y^*)dx - \gamma_1 \int_{g(y^*)}^{g^*} f(x,y^*)dx + \gamma_2 - \gamma_3 \alpha \int_0^{g(y^*)} xf(x,y^*)dx, \end{aligned} \tag{11}$$

where the second equality uses the definition of y^* in equation (5) to plug in for $\lambda b y^*$. Consider the four terms in the final line. The first two terms result from the potential discontinuity in the vaccination policy at y^*, i.e. $g(y^*) \neq g^*$. Individuals with $y = y^*$ and x between $g(y^*)$ and g^* go from being unvaccinated and not interacting – with value $-cx$ to the planner – to vaccinated and interacting – with value $-\gamma_1$ to the planner – or vice versa, depending on which threshold is higher. The final two terms are the marginal benefit of relaxing the incentive compatibility constraint and tightening the equilibrium transmission rate constraint because more individuals are below the interaction threshold and therefore join the interaction pool.

We now complete the characterization of the second-best policy in Proposition 2, which combined with Proposition 1 constitutes our main result.

Proposition 2. *Define*

$$y_0 \triangleq \frac{\gamma_1 - \gamma_3\alpha}{\lambda b}. \tag{12}$$

Then, $y_0 < y*$ *and for all* $y \leqslant y_0$*,* $g(y) = 1$*, i.e. no individuals with type* $y \leqslant y_0$ *are vaccinated. For all* $y \in (y_0, y^*]$*,*

$$g(y) = \frac{\gamma_1}{\lambda by + \gamma_3\alpha}, \tag{13}$$

which is strictly positive, strictly decreasing, and approaches 1 as y *approaches the lower limit* y_0*. The optimal vaccination policy is monotone and exhibits an exposure premium on* $[0, y^*]$ *and on* $(y^*, 1]$*.*

If $y^* < 1$*, then the vaccination threshold for* $y > y^*$ *is*

$$g^* = min\left\{\frac{\gamma_1}{c}, 1\right\}.$$

Moreover, g^* *is strictly greater than* $g(y^*)$*, the vaccination threshold at* y^**.*

The optimal policy displays three striking features. First, if $y^* < 1$, it is *non-monotone*. Specifically, we show that $g(y^*) < g^*$, which implies that the policy tends to vaccinate people with an intermediate y (just below y^*) rather than a higher or lower y. Intuitively, while high y agents efficiently self-isolate and low y agents efficiently do not, agents with intermediate y interact when they should self-isolate. They are thus particularly costly to society when unvaccinated and it is valuable to vaccinate them.

Second, the optimal policy exhibits an *exposure premium* on $\{(x, y) \mid y \leqslant y^*\}$. If the planner only took into account an individual's private value of vaccination, then the optimal policy for $y \leq y^*$ would be an xy-policy. Taking into account the spillover effect, the planner is willing to vaccinate high x individuals with relatively low risk levels, xy.

Third, as discussed previously, the policy is *invariant* to y on $\{(x, y) \mid y > y^*\}$ but does depend on x. Moreover, it exhibits an exposure premium for $y > y^*$, but for a different reason than for $y \leqslant y^*$. Below y^*, the policy exhibits an exposure premium because x matters more than y due

Covid Economics 65, 20 January 2021: 1-46

Covid Economics 65, 20 January 2021: 1-46

to spillovers. Above y^*, the policy exhibits an exposure premium because self-isolating behavior eliminates the importance of changes in vulnerability, y.

Taken together, the second and third features illustrate how to optimally allocate vaccines *within* groups as defined by behavior. On the other hand, the first feature illustrates how to allocate vaccines *between* these two groups based on the incentives that affect individuals' interaction decisions. Since individuals over interact, the optimal policy allocates more vaccines to the interacting group, generating the non-monotonicity at y^*.

Figure 2 demonstrates these features graphically.

3.3 First-Best Policy

To further understand the distortions imposed by incentives, we solve the *first-best problem*, in which the planner ignores the incentive constraint for interaction decisions. This is best thought of as a policy that jointly prescribes vaccine allocation as well as *mandatory lockdown*, but rather than apply uniformly across all individuals as is often the case in practice, the rules can be variably enforced based on observable measures of risk and sociability.[12] The planner's problem is then to choose a vaccine policy $v : [0,1]^2 \to [0,1]$, an interaction policy $\sigma : [0,1]^2 \times \{0,1\} \to [0,1]$, and a transmission rate $\lambda \in [0,1]$ to maximize the Lagrangian

$$\mathcal{L} = \mathcal{W} + \gamma_1 \left(\beta - \int_0^1 \int_0^1 v(x,y) f(x,y) dx dy \right) \\ + \gamma_3 \left(\lambda - \alpha \int_0^1 \int_0^1 (1 - v(x,y)) \sigma(x,y,0) x f(x,y) dx dy \right), \tag{14}$$

where an unvaccinated individual with type (x, y) interacts with probability $\sigma(x, y, 0)$, a vaccinated individual interacts with probability $\sigma(x, y, 1)$, and welfare is defined as before. We have the following result that shows how the planner's problem differs from before.

Proposition 3. *The planner's problem is equivalent to one in which they choose a decreasing continuous function $g : [0,1] \to [0,1]$ and real numbers $y^* \in [0,1]$ and $\lambda \in [0,1]$ to maximize the*

[12] As we show later, the lockdown rule can be implemented with a simple, constant tax on interactions.

Covid Economics 65, 20 January 2021: 1-46

Lagrangian

$$\mathcal{L} = \mathcal{W} + \gamma_1 \left(\beta - \int_0^{y^*} \int_{g(y)}^1 f(x,y)dxdy - \int_{y^*}^1 \int_{g(y^*)}^1 f(x,y)dxdy \right) + \gamma_3 \left(\lambda - \alpha \int_0^{y^*} \int_0^{g(y)} xf(x,y)dxdy \right),$$

where

$$\mathcal{W} = -\lambda b \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy - c \int_{y^*}^1 \int_0^{g(y^*)} xf(x,y)dxdy.$$

The optimal interaction threshold is given by

$$y^* = min\left\{ \frac{c - \gamma_3 \alpha}{\lambda b}, 1 \right\}$$

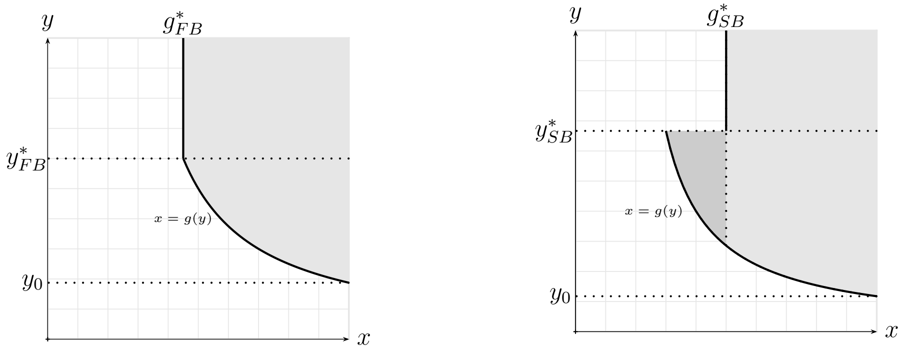
and the vaccination threshold function is

$$g(y) = min\left\{ \frac{\gamma_1}{\lambda b y + \gamma_3 \alpha}, 1 \right\},$$

which implies an optimal vaccination policy that is monotone and exhibits an exposure premium everywhere. Finally, if $y^* < 1$, then $g(y^*) = \gamma_1 / c < 1$.

The key qualitative difference between the first- and second-best vaccination policies is that the threshold function is continuous in the former case but discontinuous in the latter case at the interaction threshold, y^*. By extension, the first-best policy is monotone while the second-best policy is not. As mentioned earlier, the discontinuity at y^* in the second-best policy is driven by the fact that the threshold y^* is inefficiently high. In response, the planner reallocates vaccines toward individuals who inefficiently over-interact (the dark shaded area in Figure 2). This is no longer the case in the first-best in which y^* is chosen efficiently. On the other hand, the second and third features of the second-best policy we described earlier – constant threshold function above y^* and decreasing threshold function that implies an exposure premium below y^* – still hold; they only depend on the existence of self-isolating and interacting individuals and are not related to

Figure 2: Optimal Policies



Left panel: first-best policy. Black line $g(y) = x$. Shaded area: $v(x, y) = 1$. Right panel: second-best policy. Black line $g(y) = x$. Shaded area: $v(x, y) = 1$. Dark shaded area: additional vaccination due to incentive distortions.

the efficiency of the decision to interact.

The following result states that the first-best interaction policy can be simply implemented with a positive flat interaction tax.

Corollary 1. *Suppose in the solution to the planner's first best problem, the interaction threshold is y^* and the equilibrium transmission rate is λ. The planner can decentralize the optimal interaction policy with a flat interaction tax equal to $c - \lambda b y^*$ per interaction if $y^* < 1$ and 0 otherwise. If $y^* < 1$, then the optimal tax is strictly positive. If $y^* = 1$, then the planner's first-best problem is the same as the second-best and the optimal interaction policy is trivially decentralized.*

Weak Incentives for Self-Isolation – If the cost of interaction avoidance relative to the cost of interaction, $c/(\alpha b)$, or the fraction of the population that is vaccinated, β, is sufficiently high, then $y^* = 1$ under any policy.[13] In this case, it is as if we are in the baseline model in which all individuals always interact when given the opportunity. Second-best and first-best policies are the same since the planner sets y^* at the individually optimal level, 1, even when unconstrained. The optimal policy is monotone and exhibits an exposure premium, much as in the right panel of

[13]For example, if $c/(\alpha b) > 1$, then even an individual facing certain death upon infection would choose to interact; if β is sufficiently large then $\lambda < c/b$ for any policy that satisfies the vaccine supply constraint.

Covid Economics 65, 20 January 2021: 1-46

Figure 1.

4 Other Applications

4.1 Vaccine Hesitancy

In reality, people might be unwilling to be vaccinated. There are important qualitative differences between self-isolation and vaccine hesitancy. Crucially, in the former, *all individuals* have agency over their choices and thus outcomes, while in the latter only those who are offered a vaccine can exercise their will. This difference leads to qualitative differences in second-best policies, and thus has direct policy implications.

Two leading explanations for vaccine hesitancy exist. First, people might believe the efficacy of the vaccine to be limited. Second, people might believe that the vaccine entails unintended, negative side-effects. We focus on the latter.[14] To this end, we extend the model to include a fixed cost $p > 0$ from taking the vaccine.[15] For simplicity, in this subsection, we suppose all individuals interact when given the opportunity, i.e. $c = \infty$. A *strategy profile* is a mapping $\sigma : [0,1]^2 \to \{0,1\}$.

Simple algebra demonstrates that type (x, y) accepts the vaccine if and only if

$$\underbrace{b\lambda xy}_{\text{benefit}} \geqslant \underbrace{p}_{\text{cost}} \tag{15}$$

Equilibrium strategies are thus described by a scalar A^* such that $\sigma(x,y) = 1$ if and only if $xy \geqslant A^*$ and that solves

$$\frac{p}{bA^*}\begin{cases} \geqslant \lambda & A^* = 1 \\ = \lambda & A^* \in (0,1) \\ \leqslant \lambda & A^* = 0 \end{cases}$$

[14] The former can easily be accommodated in the baseline analysis.

[15] The cost p can be micro-founded as follows: suppose there is a hidden, aggregate state $\theta \in \{0, \bar{\theta}\}$ such that if an agent accepts vaccination, their terminal payoff is $-\theta$, and agents share a common belief $p = \mathbb{E}(\theta)$.

Covid Economics 65, 20 January 2021: 1-46

the indifference condition for the marginal individual. If the indifference condition holds with equality, then from the definition of λ, we have that

$$\frac{p}{bA^*} = \alpha \left[\int_{xy<A^*} xf(x,y)dxdy + \int_{xy \geqslant A^*} x\big(1 - v(x,y)\big) f(x,y)dxdy \right]. \tag{16}$$

Lemma 2. *For all $v \in V$, a unique v-equilibrium σ^* exists, featuring vaccination acceptance rules defined by an iso-risk threshold A^* such that a type (x,y) individual refuses vaccination if and only if $xy < A^*$. Furthermore, $A^* > 0$, i.e. a positive measure of agents would refuse vaccination under σ^*.*

One immediate prediction is that individuals with high x but very low y and vice versa are unlikely to accept vaccination. This is in line with recent surveys that find that the very old and the very young are more likely than others to refuse vaccination. That said, some surveys report that hesitancy is simply decreasing in y. The analysis could easily be adjusted to accommodate this empirical regularity by assuming that agents are offered the vaccine only after entering (or not) the interaction pool.[16]

The planner's problem is then to choose $A^* \in [0,1]$, $\lambda \in [0,1]$, and a vaccination policy $v : [0,1]^2 \to [0,1]$ to maximize the Lagrangian

$$\begin{aligned} \mathcal{L} = \mathcal{W} + \gamma_1 \left(\beta - \int_0^1 \int_0^1 v(x,y) f(x,y)dxdy \right) + \gamma_2 \left(A^* - \min\left\{ \frac{p}{\lambda b}, 1 \right\} \right) \\ + \gamma_3 \left(\lambda - \alpha \left[\int_{(x,y):xy \geqslant A^*} x(1 - v(x,y)) f(x,y)dxdy + \int_{(x,y):xy<A^*} xf(x,y)dxdy \right] \right), \end{aligned} \tag{17}$$

where welfare is

$$\mathcal{W} = -\int_{(x,y):xy \geqslant A^*} \left((1 - v(x,y))\lambda byx + v(x,y)p \right) f(x,y)dxdy - \int_{(x,y):xy<A^*} \lambda byxf(x,y)dxdy.$$

Proposition 4. *The second-best policy is an x-threshold policy characterized by an x-threshold*

[16] Specifically, the optimal policy would consist of a binding $y = A^*$ curve, rather than an $xy = A^*$ curve, with the remaining construction of $g(y)$ virtually unchanged.

Covid Economics 65, 20 January 2021: 1-46

function $g : [0,1] \to [0,1]$, *i.e.*

$$v(x,y)\sigma(x,y) = \begin{cases} 0 & x < g(y) \\ 1 & x \geqslant g(y) \end{cases}$$

Define

$$\hat{y} \triangleq \max\left\{0, \frac{\gamma_1 - \gamma_3\alpha + p}{b\lambda}\right\} \tag{18}$$

and

$$\bar{y} \triangleq \frac{\gamma_3 \alpha A^*}{\gamma_1}. \tag{19}$$

The threshold function g is given by

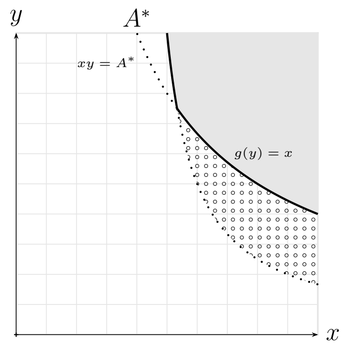
$$g(y) = \begin{cases} 0 & y < \hat{y} \\ \frac{\gamma_1 + p}{\lambda b y + \gamma_3 \alpha} & y \in [\hat{y}, \bar{y}) \\ A^*/y & y \geqslant max\{\hat{y}, \bar{y}\} \end{cases}$$

The optimal policy is monotone, is an xy-policy with iso-risk threshold A^* for all $y \in [\hat{y}, \bar{y})$, and exhibits an exposure premium for all $y \geqslant max\{\hat{y}, \bar{y}\}$.

The proposition shows that for $y \geqslant \bar{y}$, vaccine hesitancy is irrelevant because all those offered a vaccine strictly prefer to accept. Yet, for $y \in [\hat{y}, \bar{y})$, vaccine hesitancy is a binding constraint and the vaccine is simply offered to all who will take it. This is because the optimal policy exhibits an exposure premium, while vaccine hesitancy depends only on the unconditional risk level, xy. The marginal individual vaccinated with a high y has a relatively low x, which implies a low spillover and therefore a relatively high risk level, xy. The opposite is true for the marginal individual vaccinated with a low y. Figure 3 demonstrates the policy graphically, showing explicitly how inefficient vaccine rejection occurs by those with high x and low y.

Covid Economics 65, 20 January 2021: 1-46

Figure 3: Second-best policies under vaccine hesistancy and one-way efficacy



Left panel: second-best under vaccine hesitancy. Black line: optimal policy threshold. Dotted area: inefficient vaccine rejection. Grey shaded area: $v(x, y) = 1$. Right panel: second-best under one-way efficacy, xy-policy. Grey shaded area: $v(x, y) = 1$.

4.2 Silent-spreading

Thus far, we have assumed that a vaccinated agent cannot contract the disease in the first place, and hence can neither experience severe symptoms nor transmit the illness to others, i.e. that the vaccine is *two-way*. As of the writing of this paper, it is not clear whether this assumption will hold in the case of COVID-19 vaccines. Of the available vaccines, only the Oxford-AstraZeneca vaccine ran bi-weekly PCR tests that indicate the assumption is valid.[17] Furthermore, some experts remain uncertain that the vaccines developed are two-way.[18] We may analyse a *one-way* vaccine by supposing that vaccination sends an individual's y to 0, but leaves the possibility of transmission unaffected. That is, a vaccinated individual is protected, but interaction with them entails unmitigated risk for unvaccinated agents. Again assuming full interaction ($c = \infty$) for simplicity, the optimal vaccine policy in such a case would be an xy-policy (that does *not* exhibit an exposure premium), as the value of vaccination is purely determined by an agent's risk, xy. Taking this and the main result together, it is clearly imperative that the two-way feature of a

[17]See https://tinyurl.com/y8n2ayrh for a detailed discussion.

[18]Professor Robert Read states: "*The reason [not to vaccinate young people] is that a) they don't get such a severe disease and b) we haven't been able to demonstrate yet that the vaccines have any impact at all on transmission.*" See https://tinyurl.com/y5d7ucr6.

CEPR PRESS

Covid Economics 65, 20 January 2021: 1-46

vaccine is carefully verified.

5 Practical Guidelines

Our analysis offers clear, practical guidelines to government health agencies tasked with developing vaccine prioritization schemes. How are countries approaching this problem in reality? There is wide-spread consensus that front-line health workers – who arguably rank highly on both fronts – should be among the very first to be vaccinated.[19] Beyond this consensus, prioritization approaches tend to differ, sometimes dramatically, between countries. For instance, most Western countries such as the US, UK, Canada and Europe are vaccinating based on y, whereas countries such as Indonesia and Russia are vaccinating the youth. In the case of the former, Amin Soebandrio, director at the Eijkman Institute for Molecular Biology in Jakarta, states:[20]

> "*Our aim is herd immunity…with the most active and exposed group of population – those 18 to 59 – vaccinated, they form a fortress to protect the other groups. It's less effective when we use our limited number of vaccines on the elderly when they're less exposed.*"

These comments raise another interesting point of contention, namely that the optimal vaccination policy should be carefully calibrated to the limitations on supply, and that with very limited supply, priority should be re-balanced toward high x agents, rather than high y. In contrast, European Union guidelines suggest the opposite: when supply is limited, priority should be given to high y people.[21]

With these observations in hand, we perform a series of numerical calculations, not only to uncover deeper properties of the second-best policy not amenable to analytic derivation, but also to compare it to commonly considered heuristics as alluded to above. These include: 1) y-policies, which vaccinate only the most vulnerable; 2) x-policies, which vaccinate only the most interactive;

[19] Healthcare workers also ranked highly for other reasons. Most obviously, they reduce overall mortality rates. Beyond this, they are often given priority as a way of acknowledging their sacrifice and social service prior to the vaccine existing. See Pathak et al. (2020) for a discussion.

[20] See https://tinyurl.com/y8vvpfkm and https://tinyurl.com/yybj498g.

[21] See https://tinyurl.com/y8n2ayrh.

Covid Economics 65, 20 January 2021: 1-46

3) xy-policies, which vaccinate those who face the most risk overall; and 4) random policies under which vaccinations are uncorrelated with observable characteristics. Throughout these exercises, we assume x and y are drawn independently from standard uniform distributions, unless otherwise stated.[22]

Before proceeding, we show that other than the choice of distribution F, there are essentially only two parameters in the model; the one we discuss shortly and β. As such, since we vary both in our numerical exercises, they show the full range of model outcomes across the entire parameter space taking as given our choice for F.

Particularly for the final set of figures, we define a key invariant of the model – the *survival value index (SVI)* as $\alpha b/c$ – which is essentially the value of life, normalized by the cost of remaining infection-free, i.e. self-isolating. Crucially, welfare comparisons depend only upon this invariant, rather than on each of its deep parameters separately.[23]

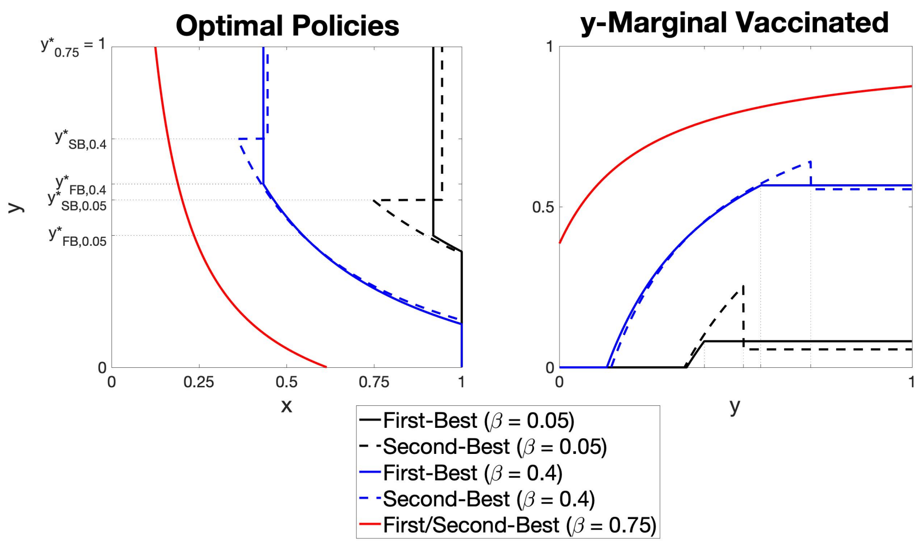
Lemma 3. *The optimal vaccination policy and equilibrium welfare normalized by c do not depend on α, b, or c independently, but only on the survival value index (SVI), $\alpha b/c$.*

We begin with a series of graphs comparing the first- and second-best policies. Figure 4 shows the basic features of the first- and second-best policies. Note in the panel entitled "Optimal Policies" that the first-best policies demonstrate a greater exposure premium than the second-best; spillovers are lessened in the second-best problem because behavior, y^*, responds to changes in the environment. The "y-Marginal Vaccinated" panel plots the y-marginal conditional on vaccination, showing clearly the non-monotonicity of the second-best policy – intermediate y agents constitute the greatest proportion of vaccinated agents, whereas the first-best is monotone. Figure 5 compares the performance of the first- and second-best policies. The panel entitled "Incentive Distortion" measures the performance of the second-best – without mandatory lockdown policy – relative to the first-best. In the "Policy Distortion" panel, we measure the impact that ignoring behavior

[22]As such, we avoid drawing inferences regarding magnitudes, focusing on qualitative insights.

[23]Note that changes in α, b, or c, holding $\alpha b/c$ fixed, can have effects on observable variables. For example, a higher b but lower α implies a higher utility loss from death, but a lower risk of death contingent on interaction. The planner and individuals do not care about the difference, but the lower α case has a lower observed death rate.

Figure 4: Optimal Policies: Basics

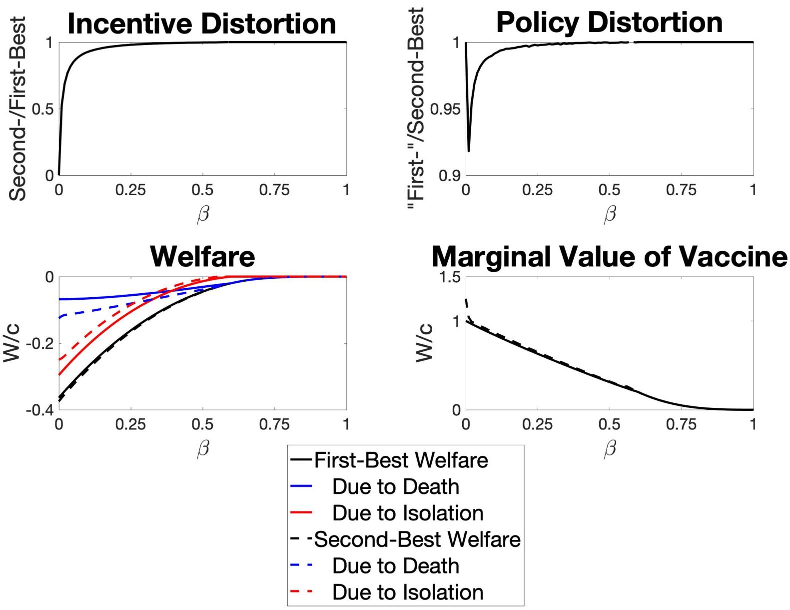


Left panel: x-threshold functions under the first- and second-best policies at various levels of the percentage of the population that can be vaccinated, β. Individuals to the right of the appropriate curve are vaccinated. Right panel: the marginal distributions of y conditional on being vaccinated under the first- and second-best policies at various levels of β. Both panels reflect that $\beta = 0.75$ is sufficiently large so that the first- and second-best policies are the same.

has on welfare by studying a novel policy wherein the planner solves the second-best problem, but assuming that agents internalize externalities and choose y^* efficiently. Thus, the planner sets policy erroneously believing that agents self-isolate according to the optimal mandatory lockdown (first-best) rule. These first two panels show that the distinction between the first- and second-best (and policymaker awareness of which holds) is particularly important when vaccine supply is low. The panel entitled "Welfare" decomposes welfare losses into losses from deaths and losses from self-isolation.[24] First-best losses are driven primarily by isolation rather than death, reflecting that private incentives to self-isolate are inefficiently weak. Finally, the "Marginal Value of Vaccine" panel demonstrates both that there are diminishing returns to vaccination, and that the marginal value is greater under the second-best. This implies that vaccine supply should be geographically

[24]Throughout this section, welfare is normalized by c so that it only depends on $\alpha b/c$, not on each parameter separately.

Figure 5: Optimal Policies: Analysis



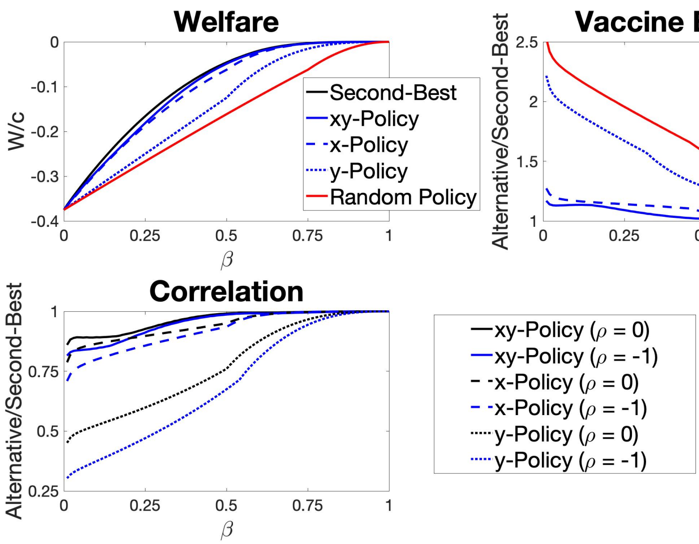
Top-left panel: the improvement in welfare (compared to no vaccines) under the second-best relative to under the first-best at various levels of the percentage of the population that can be vaccinated, β. Top-right panel: the improvement in welfare (compared to no vaccines) under the misperceived first-best relative to under the second-best at various levels of β. The misperceived first-best is derived by optimizing assuming individuals will self-isolate efficiently as in the first-best even though they self-isolate selfishly as in the second-best. Bottom-left panel: welfare (normalized by c) and its components under the first- and second-best policies at various levels of β. Bottom-right panel: the marginal normalized welfare value of a vaccine under the first- and second-best at various levels of β.

dispersed, not concentrated, and should favor regions that cannot implement mandatory lockdown policies.

We turn next to heuristic policies, wherein we compare the second-best against x-, y-, xy-, and random policies. Figure 6 compares welfare across these policies by looking at: 1) absolute welfare in "Welfare"; 2) a novel performance measure that shows the relative quantity of vaccines required so that each heuristic policy performs as well as the second-best in "Vaccine Equivalent"; and 3) in "Correlation", the welfare improvement (compared to no vaccines) under heuristic policies relative to under the second-best for the baseline type distribution and an alternative distribution with

Covid Economics 65, 20 January 2021: 1-46

Figure 6: Heuristic Policies: Welfare



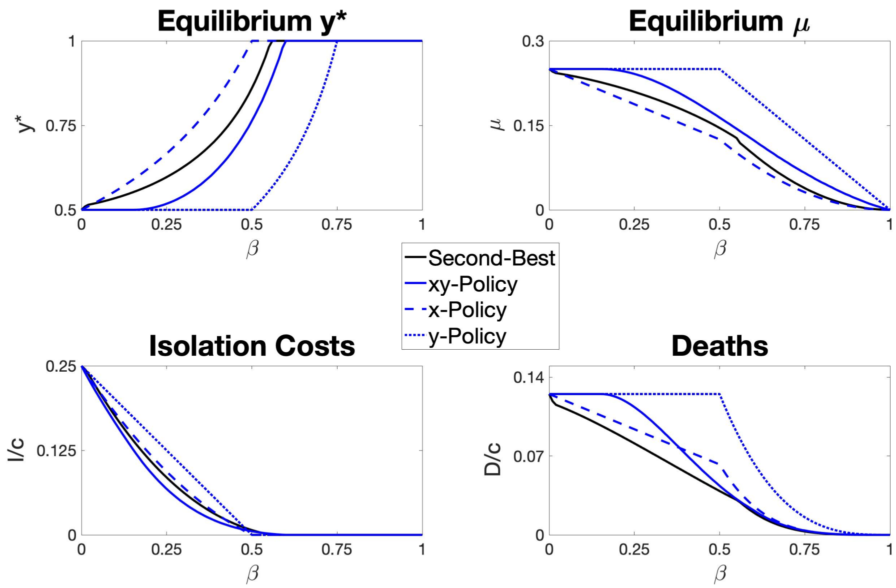
Top-left panel: welfare (normalized by c) under the second-best and heuristic policies. Top-right panel: the multiple of vaccines required so that each heuristic policy delivers the same welfare as the second-best. For example, a value of 1.5 means a 50% increase in vaccines is required. Bottom-left panel: the improvement in welfare (compared to no vaccines) under each heuristic policy relative to the improvement under the second-best for the baseline distribution with independent uniforms ($\rho = 0$) and an alternative distribution with uniform marginals but a negative correlation between x and y ($\rho = -1$.

uniform marginals but a negative correlation between x and y.[25] The most striking observation is that y-policies perform significantly worse than other heuristics, and that this performance loss is greatest when supply is limited. This is of significant practical relevance, as the majority of countries not only adopt such policies, but do so at the start of vaccine roll-out when supply is most constrained. To better understand this result, consider Figure 7, which shows equilibrium interaction thresholds, y^*, interaction pool sizes normalized by α, μ, and the components of welfare losses under the second-best and heuristic policies. A policy reduces isolation costs through direct and indirect channels: allowing those vaccinated to interact freely and encouraging interaction by reducing the size of the *unvaccinated* interaction pool. y-policies work through the direct channel,

[25]Specifically, we impose the following functional form for the pdf of the distribution: $f(x, y) = 1+\rho(1-2x)(1-2y)$, where $\rho \in [-1, 1]$. A little algebra shows that the marginal distributions of x and y are uniform, that the conditional distributions have affine pdfs, and that the correlation between x and y is $\rho/3$.

Covid Economics 65, 20 January 2021: 1-46

Figure 7: Heuristic Policies: Details



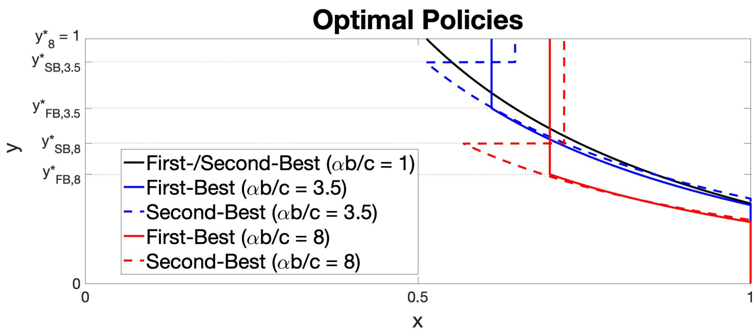
Top panels: the equilibrium interaction threshold, y^*, and unvaccinated interaction pool size, μ, under the second-best and heuristic policies at various levels of the percentage of the population that can be vaccinated, β. Bottom panels: the components of welfare losses (normalized by c), isolation costs or death, under the second-best and heuristic policies at various levels of β. Here, an improvement in welfare is a fall in isolation costs or deaths.

but shut the indirect channel down; targeting the self-isolating leaves the interaction threshold and unvaccinated interaction pool unchanged. Crucially, when vaccine supplies, β, are low and many are self-isolating, the indirect channel is of first-order importance. Policies that target high x individuals thus perform far better, even at reducing isolation costs among the most vulnerable. The "Correlation" panel shows that this issue is exacerbated when x and y are negatively correlated because the y-policy ends up targeting low x individuals who are not particularly valuable to vaccinate. Finally, the novel measure in "Vaccine Equivalent" captures the trade-off between implementing optimal policy – which might require costly or delayed implementation – or using well-established heuristics that are easier to roll-out.

In each of the last three figures, we vary the SVI and show how key outcomes and measure change. In Figure 8, we show the x-threshold functions under first- and second-best policies. In

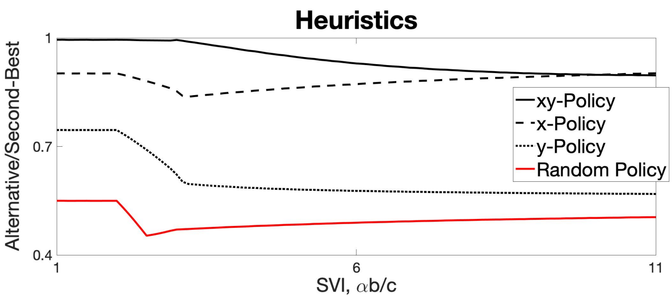
Covid Economics 65, 20 January 2021: 1-46

Figure 8: Infectiousness: Basics



x-threshold functions under the first- and second-best policies at various levels of $\alpha b/c$ (SVI). Individuals to the right of the appropriate curve are vaccinated. An SVI of 8 is sufficiently large so that the first- and second-best policies are the same.

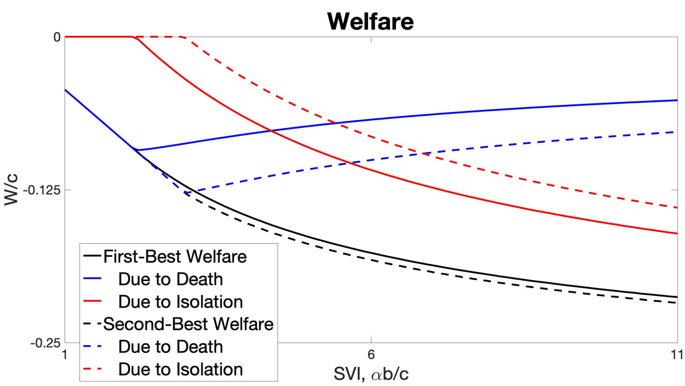
Figure 9: Infectiousness: Heuristics



The improvement in welfare (compared to no vaccines) under each heuristic policy relative to the improvement under the second-best at various levels of $\alpha b/c$ (SVI).

Figure 9, we show the improvement in welfare (compared to no vaccines) under heuristic policies relative to under the second-best. In Figure 10, we show normalized welfare and its components under the first- and second-best policies. It is perhaps easiest to view an increase in the SVI as being driven by α, holding b and c fixed. It then corresponds to an increase in the infectiousness of the virus, holding individual-specific fixed risk, vaccine supply, the value of life, and the cost of self-isolation. As such, these exercises speak directly to current concerns regarding new variants of COVID-19 that exhibit these features. An important takeaway is that as infectiousness, α, increases, policy must take into account the response of behavior, y^*. More individuals find interaction too costly and instead self-isolate. It is valuable to vaccinate such individuals, but recall that the differences in y among them do not matter since they no longer face any risk of infection. As such, the optimal policy spreads vaccines across a large range of y, targeting only those with the highest x. We can see this result in Figure 8 as well as in Figure 9, which shows that y-policies perform particularly poorly as infectiousness rises. Finally, in Figure 10, we can see that individual behavior and optimal policy respond so aggressively to the increase in infectiousness that the welfare loss from death actually falls while interaction avoidance costs rise.

Figure 10: Infectiousness: Welfare



Welfare (normalized by c) and its components under first- and second-best policies at various levels of $\alpha b/c$ (SVI).

References

Babus, A., S. Das, and S. Lee, "The Optimal Allocation of Covid-19 Vaccines," *Covid Economics*, 2020, *44*.

Ballester, C., A. Calvó-Armengol, and Y. Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 2006, *74* (5), 1403–1417.

Brotherhood, L., P. Kircher, C. Santos, and M. Tertilt, "An economic model of the Covid-19 pandemic with young and old agents: Behavior, testing and policies," *working paper*, 2020.

Chen, F. and F. Toxvaerd, "The economics of vaccination," *Journal of Theoretical Biology*, 2014, *363*, 105–117.

Epstein, J.M., J. Parker, D. Cummings, and R.A. Hammond, "Coupled contagion dynamics of fear and disease: mathematical and computational explorations," *PLoS ONE*, 2008, *3* (12).

Farboodi, M., G. Jarosch, and R. Shimer, "The Emergence of Market Structure," *working paper*, 2020.

Funk, S., M. Salathé, and V.A.A. Jansen, "Modelling the influence of human behaviour on the spread of infectious diseases: a review," *Journal of the Royal Society*, 2010, *7* (50).

Galeotti, A., B. Golub, and S. Goyal, "Targetting Interventions in Networks," *Econometrica*, 2020, *88* (6), 2445–2471.

Geoffard, P.-Y. and T. Philipson, "Rational Epidemics and Their Public Control," *International Economic Review*, 1996, *37* (3).

Greenwood, J., P. Kircher, C.Santos, and M. Tertilt, "An Equilibrium Model of the African HIV/AIDS Epidemic," *Econometrica*, 2019, *87* (4).

Gross, T., C.J.D D'Lima, and B. Blasius, "Epidemic dynamics on an adaptive network," *Phys Rev Lett*, 2006, *96* (20).

Covid Economics 65, 20 January 2021: 1-46

Kremer, M., "Integrating behavioral choice into epidemiological models of AIDS," *Quarterly Journal of Economics*, 1996, *111*, 549–573.

Pastor-Satorras, R. and A. Vespignani, "Immunization of complex networks," *Phys Rev E*, 2002, *65* (3).

Pathak, P.A., T. Sönmez, M.U. Ünver, and M.B. Yenmez, "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing," *working paper*, 2020.

Rowthorn, R. and F. Toxvaerd, "On the Management of Population Immunity," *working paper*, 2020.

Shaw, L.B. and I.B. Schwartz, "Fluctuating epidemics on adaptive networks," *Phys Rev E*, 2008, *77* (6).

Talamás, E. and R. Vohra, "Free and perfectly safe but only partially effective vaccines can harm everyone," *Games and Economic Behaviour*, 2020, *122*.

Covid Economics 65, 20 January 2021: 1-46

A Proofs

A.1 Lemma 1

Recall that the threshold is $y^* = c/(\lambda b)$, which is decreasing in λ. It follows that the right-hand-side of equation 4 is decreasing in λ since the integrand is always positive and as λ increases, y^* decreases, and the integral is over a smaller set. The right-hand-side is also continuous by the Fundamental Theorem of Calculus. Furthermore, it is strictly positive when $\lambda = 0$ and is always strictly less than 1. Since the left-hand-side is λ, the result follows from the Intermediate Value Theorem.

A.2 Proposition 1

First, we have the following lemma that shows that the planner always strictly prefers more vaccines.

Lemma A.1. *The Lagrange multiplier on the vaccine supply constraint, γ_1, is strictly positive.*

Proof. If there is a strictly positive measure of self-isolating individuals, then they can be vaccinated for a strictly positive private benefit without any spillovers. As such, the marginal benefit of a vaccine must be strictly positive. If there is not a strictly positive measure of self-isolating individuals, then since $\beta < 1$, there is a strictly positive measure of unvaccinated interacting individuals. They can be vaccinated for a strictly positive private benefit with spillovers that yield a strictly positive public benefit. To see this note that as the equilibrium interaction rate falls, other unvaccinated interacting individuals' infection rates fall, which raises welfare, and at most a measure 0 of isolating individuals can begin interacting since there are at most a measure 0 of isolating individuals. □

We now prove the following proposition that is also useful for proving Proposition 2.

Proposition A.1. *The optimal policy is characterized by a function $g : [0, 1] \to [0, 1]$ and a number $\bar{g}$ such that*

$$v(x,y) = \begin{cases} 0 & x \le g(y); y \le y^* \\ 1 & x > g(y); y \le y^* \\ 0 & x \le g^*; y > y^* \\ 1 & x > g^*; y > y^* \end{cases}$$

i.e. a type (x, y) individual with $y \le y^*$ is vaccinated if and only if $x > g(y)$ and an individual with $y > y^*$ is vaccinated if and only if $x > g^*$. Moreover, the threshold function below y^* is given by

$$g(y) = min\left\{\frac{\gamma_1}{\lambda b y + \gamma_3 \alpha}, 1\right\} \tag{A.1}$$

and the threshold above y^* is given by

$$g^* = min\left\{\frac{\gamma_1}{c}, 1\right\}. \tag{A.2}$$

Proof. First, consider type (x, y) with $y > y^*$. The first derivative of the planner's Lagrangian in equation (7) with respect to $v(x, y)$ is

$$(cx - \gamma_1)f(x, y),$$

where the first term is the benefit that the individual no longer has to pay the interaction avoidance cost and the second term is the cost of providing the vaccine. Observe that the second derivative is always 0. If $\gamma_1/c > 1$, then since $x \leq 1$, the planner's derivative is always strictly negative and they optimally set $v(x, y) = 0$. If $\gamma_1/c = 1$, then the planner's derivative is strictly negative if $x < 1$, but always equal to 0 if $x = 1$. Since $x = 1$ with probability 0, we set $v(x, y) = 0$ regardless. If $\gamma_1/c < 1$, then the planner's derivative is always strictly negative if $x < \gamma_1/c$, equal to 0 if $x = \gamma_1/c$, and strictly positive if $x > \gamma_1/c$. As such, from the definition of $\bar{g}$ in equation (A.2), the planner optimally sets $v(x, y) = 1$ if and only if $x > \bar{g}$.

Next, consider type (x, y) with $y \leq y^*$. In this case, the first derivative of the planner's Lagrangian with respect to $v(x, y)$ is

$$(\lambda byx - \gamma_1 + \gamma_3\alpha x)f(x, y), \tag{A.3}$$

where the first term is the benefit that the individual can no longer become infected, the second term is the cost of providing the vaccine, and the third term captures the externality of lowering the equilibrium transmission rate by vaccinating an individual who was interacting when unvaccinated. Again, observe that the second derivative is always 0. Suppose $\lambda by + \gamma_3\alpha \geq \gamma_1$. It follows that the derivative in equation (A.3) is strictly increasing in x. Moreover, the derivative is strictly negative if $x < \gamma_1/(\lambda by + \gamma_3\alpha)$, strictly positive if the opposite strict inequality holds, and equal to 0 if the inequality holds with equality. It follows from the definition of $g(\cdot)$ in equation (A.1) that the planner optimally sets $v(x, y) = 1$ if and only if $x > g(y)$.

Finally, consider type (x, y) with $y \leq y^*$ and $\lambda by + \gamma_3\alpha < \gamma_1$. If $\lambda by + \gamma_3\alpha \geq 0$, then the derivative in equation (A.3) is weakly increasing in x and is strictly negative if $x = 1$, and so is strictly negative for any x. If $\lambda by + \gamma_3\alpha < 0$, then the derivative is strictly decreasing in x. Since the derivative is strictly negative if $x = 0$, it follows that the derivative is strictly negative for any x. As such, in any case, the planner optimally sets $v(x, y) = 0$. This is consistent with $g(y)$ because $\gamma_1/(\lambda by + \gamma_3\alpha) > 1$, so $g(y) = 1$. □

We immediately have the following result.

Corollary 2. *The threshold function g is continuous and decreasing.*

Proposition 1 follows from plugging in the vaccination policy

$$v(x, y) = \begin{cases} 0 & x \leq g(y); y \leq y^* \\ 1 & x > g(y); y \leq y^* \\ 0 & x \leq g^*; y > y^* \\ 1 & x > g^*; y > y^* \end{cases}$$

A.3 Proposition 2

Following Proposition A.1, all that remains to prove in the first part of Proposition 2 is that $y_0 < y^*$, $g(y)$ is strictly positive for all $y \leq y^*$, and $g(\cdot)$ exhibits an exposure premium $[0, y^*]$ and on $(y^*, 1]$.

We begin by proving a lemma necessary for further characterizing the second best because it shows how to deal with the case that $\lambda = c/b$. Call Problem 1 the planner's problem in which we assume $y^* < 1$ and given by $y^* = c/(\lambda b)$. As such, Problem 1 is only defined for $\lambda > c/b$. Call Problem 2 the planner's problem in which we assume $y^* = 1$. Even though this is only the case for $\lambda \leq c/b$, define Problem 2 and welfare for all policies $g(\cdot)$ and λ, assuming that y^* is fixed at 1. We say that a potential solution to Problem 2 is feasible if $\lambda \leq c/b$.

Lemma A.2. *Let P_1 be the set of solutions to Problem 1 and P_2 be the set of solutions to Problem 2 with $\lambda \leq c/b$. The optimal policy is the choice from the union of these sets that maximizes welfare.*

Proof. Observe that Problem 2 always has a solution. First, we show that if a solution to Problem 2 has $\lambda > c/b$, then there is a strictly better solution to Problem 1 and it is the optimal policy. Suppose Problem 2 has such an infeasible solution. The welfare in this solution is weakly higher than the welfare from any feasible solution with $y^* = 1$, i.e. with $\lambda \leq c/b$. Now, suppose the planner implements the same policy in the true planner's problem that sets $y^* = \min\{c/(\lambda b), 1\}$ – by the same policy, we mean vaccinating the same set of types (x, y), so the vaccination policy is unaffected by changes in λ and y^*. We can see that the policy leads to $\lambda > c/b$ and $y^* < 1$. If not, then $y^* = 1$ and as in Problem 2, we must have $\lambda > c/b$, which implies $y^* < 1$, a contradiction. Moreover, the policy yields higher welfare than in Problem 2. To see this note that there are two effects on welfare from reducing y^* below 1 while holding the policy fixed: the direct effect of the change in y^* and the indirect effect through the implied change in λ. For the indirect effect, as y^* falls below 1, λ falls as well, which strictly increases welfare. For the direct effect, the partial derivative of welfare with respect to y^* is $-\lambda b y^* g(y^*) + c g(y^*)$, which is weakly negative if $y^* > c/(\lambda b)$. It follows that as y^* decreases from 1 until the incentive compatibility constraint, $y^* = c/(\lambda b)$, is satisfied, welfare increases. As such, we have a policy that generates $y^* < 1$ that yields higher welfare than any feasible policy that generates $y^* = 1$. It follows that the optimal policy leads to $y^* < 1$ and so is a well-defined solution to Problem 1.

Next, suppose no solutions to Problem 2 have $\lambda > c/b$. It follows that any solutions to Problem 2 (there must be at least one) are feasible and yield higher welfare than any policy that generates $y^* = 1$. If Problem 1 has a solution, then it yields higher welfare than any policy that generates $y^* < 1$. It follows that the optimal policy is the welfare-maximizing choice from the set of solutions to Problem 1 and from the set of solutions to Problem 2 with $\lambda \leq c/b$. If, on the other hand, Problem 1 does not have a solution, then for any policy that yields $y^* < 1$, there is another policy that yields $y^* < 1$ and strictly higher welfare. As such, the optimal policy has $y^* = 1$. It follows that the optimal policy is a feasible solution to Problem 2 and all feasible solutions to Problem 2 yield the same welfare as the optimal policy. □

A takeaway from Lemma A.2 is that if the optimal policy has $y^* = 1$, then it is a solution to the planner's problem assuming y^* is fixed at 1, in which case we can ignore the planner's choice of y^* and the derivative of the Lagrangian with respect to λ is the same as when $\lambda < c/b$:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -b \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy + \gamma_3.$$

We next show that the planner never vaccinates individuals with $x = 0$.

Lemma A.3. *For all $y \leq y^*$, the threshold function, $g(y)$, is strictly positive.*

Proof. It follows from equation (8) that for all $y \leq y^*$, if $g(y) = 0$, then the derivative of the Lagrangian with respect to $g(y)$ is strictly positive. □

We can now show that the remaining Lagrangian multipliers are strictly positive, i.e. that the planner strictly prefers to reduce y^* and interaction relative to the individually optimal level and to reduce the transmission rate λ relative to the equilibrium level.

Lemma A.4. *The Lagrange multiplier on the equilibrium transmission rate constraint, γ_3, is strictly positive. If $y^* < 1$, then the Lagrange multiplier on the incentive compatibility constraint for y^*, γ_2, is strictly positive.*

Proof. First, suppose $y^* = 1$. Lemma A.2 shows that the First Order Condition for λ is

$$\gamma_3 = b \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy.$$

Since $y^* > 0$ and Lemma A.3 shows that $g(y) > 0$ for all $y \leq y^*$, it follows that $\gamma_3 > 0$.

Next, suppose $y^* < 1$. Plugging in the definition of y^* from equation (5) into the derivative of the Lagrangian with respect to $g(y^*)$ yields

$$\frac{\partial \mathcal{L}}{\partial g(y^*)} = (\gamma_1 - (c + \gamma_3\alpha)g(y^*))f(g(y^*), y^*). \tag{A.4}$$

Suppose $\gamma_1 = (c + \gamma_3\alpha)g(y^*)$. Plugging in for c in the final line of the derivative of the Lagrangian with respect to y^* yields

$$\frac{\partial \mathcal{L}}{\partial y^*} = \left(\frac{\gamma_1}{g(y^*)} - \gamma_3\alpha\right) \int_{g(y^*)}^{g^*} xf(x,y^*)dx - \gamma_1 \int_{g(y^*)}^{g^*} f(x,y^*)dx + \gamma_2 - \gamma_3\alpha \int_0^{g(y^*)} xf(x,y^*)dx.$$

If $g^* \geq g(y^*)$, then for all $x \in [g(y^*), g^*]$, $x/g(y^*) \geq 1$. It follows that

$$\frac{\gamma_1}{g(y^*)} \int_{g(y^*)}^{g^*} xf(x,y^*)dx \geqslant \gamma_1 \int_{g(y^*)}^{g^*} f(x,y^*)dx.$$

Covid Economics 65, 20 January 2021: 1-46

If $g^* < g(y^*)$, then for all $x \in [g^*, g(y^*)]$, $x/g(y^*) < 1$. In that case, it follows that

$$\begin{aligned}\frac{\gamma_1}{g(y^*)}\int_{g(y^*)}^{g^*} xf(x,y^*)dx &= -\frac{\gamma_1}{g(y^*)}\int_{g^*}^{g(y^*)} xf(x,y^*)dx\\ &> -\gamma_1\int_{g^*}^{g(y^*)} f(x,y^*)dx\\ &= \gamma_1\int_{g(y^*)}^{g^*} f(x,y^*)dx.\end{aligned}$$

In either case, we have that

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial y^*} &\geqslant -\gamma_3\alpha\int_{g(y^*)}^{g^*} xf(x,y^*)dx + \gamma_2 - \gamma_3\alpha\int_0^{g(y^*)} xf(x,y^*)dx\\ &= \gamma_2 - \gamma_3\alpha\int_0^{g^*} xf(x,y^*)dx.\end{aligned}$$

Since the derivative of $\mathcal{L}$ with respect to y^* must be equal to 0, it follows that if γ_3 is weakly negative, then γ_2 must be weakly negative as well. Setting the derivative of $\mathcal{L}$ with respect to λ in equation (10) equal to 0 shows that one of γ_2 and γ_3 must be strictly positive. It follows that $\gamma_3 > 0$. Finally, regardless of whether $g^* \geqslant g(y^*)$ or the opposite holds, the sum of the first two terms in the last line of equation (11) for the derivative of $\mathcal{L}$ with respect to y^* are negative, which implies that

$$\frac{\partial\mathcal{L}}{\partial y^*} \leqslant \gamma_2 - \gamma_3\alpha\int_0^{g(y^*)} xf(x,y^*)dx.$$

Since $\gamma_3 > 0$ and $g(y^*) > 0$, it follows that $\gamma_2 > 0$ as well.

Next, suppose $\gamma_1 > (c + \gamma_3\alpha)g(y^*)$. It follows from equation (A.4) that the derivative of $\mathcal{L}$ with respect to $g(y^*)$ is strictly positive, so $g(y^*) = 1$. It follows that $-c > \gamma_3\alpha - \gamma_1$ and, recalling that $g^* \leqslant 1$, we have that

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial y^*} &= -c\int_{g^*}^1 xf(x,y^*)dx + \gamma_1\int_{g^*}^1 f(x,y^*)dx + \gamma_2 - \gamma_3\alpha\int_0^1 xf(x,y^*)dx\\ &> (\gamma_3\alpha - \gamma_1)\int_{g^*}^1 xf(x,y^*)dx + \gamma_1\int_{g^*}^1 f(x,y^*)dx + \gamma_2 - \gamma_3\alpha\int_0^1 xf(x,y^*)dx\\ &= \gamma_2 - \gamma_3\alpha\int_0^{g^*} xf(x,y^*)dx.\end{aligned}$$

The rest of the argument follows as in the previous paragraph.

Finally, we cannot have $\gamma_1 < (c + \gamma_3\alpha)g(y^*)$ because in that case, the derivative of $\mathcal{L}$ with respect to $g(y^*)$ is strictly negative. It then must be that $g(y^*) = 0$, which is not the case. $\square$

We can now complete the proof of the first half of Proposition 2 with the following two corollaries.

Corollary 3. *The threshold below which $g(y) = 1$, y_0, is strictly below y^*.*

Covid Economics 65, 20 January 2021: 1-46

Proof. Since $g(y_0) = 1$ and $g(\cdot)$ is decreasing, it is sufficient to show that $g(y^*) < 1$. Plugging in the definition y^* from equation (5), we have from equation (A.1) in Proposition A.1 that

$$g(y^*) = \min\left\{\frac{\gamma_1}{c + \gamma_3\alpha}, 1\right\}.$$

Since γ_1 and γ_3 are strictly positive, it follows that $g(y^*) < \gamma_1/c$. As such if $g(y^*) = 1$, then $\gamma_1/c > 1$. If $y^* < 1$, then it follows from equation (A.2) in Proposition A.1 that $g^* = 1$. Since $g(\cdot)$ is decreasing, we then have that no individuals are vaccinated, which violates the vaccine supply constraint and so cannot be the case. On the other hand, if $y^* = 1$, then since $g(\cdot)$ is decreasing, we again have that no individuals are vaccinated, which again cannot be the case. □

Corollary 4. *For all $y \leqslant y^*$, the optimal vaccination policy implied by $g(\cdot)$ is monotone and exhibits an exposure premium. For all $y > y^*$, the optimal vaccination policy is monotone and exhibits an exposure premium.*

Proof. To see that the optimal policy is monotone on $y \leqslant y^*$, note that the threshold function is decreasing. Take an (x, y) such that $y \leqslant y^*$ and $v(x, y) = 1$. If $x' \geqslant x$ and $y' \in [y, y^*]$, then $g(y') \leqslant g(y) \leqslant x \leqslant x'$, so $v(x', y') = 1$ as well.

To see that the optimal policy exhibits an exposure premium on $y \leqslant y^*$, use the optimal $g(\cdot)$ from equation (A.1) in Proposition A.1 to take the derivative of $yg(y)$ with respect to y. If $y < y_0$, then the derivative is 1. If $y \in [y_0, y^*]$, then

$$\frac{\partial yg(y)}{\partial y} = \left(1 - \frac{\lambda by}{\lambda by + \gamma_3\alpha}\right) g(y),$$

which is strictly positive since $\gamma_3 > 0$.

To see that the optimal policy is monotone and exhibits an exposure premium for $y > y^*$, note that it is an x-policy and that since the threshold is constant, the derivative of y times the threshold is 1, which is positive. □

Now, following Proposition A.1, all that remains to prove in the second half of Proposition 2 is that if $y^* < 1$, then $g^* > g(y^*)$.

Corollary 5. *If $y^* < 1$, then $g^* > g(y^*)$.*

Proof. Suppose $y^* < 1$. If $g^* = 1$, then since Lemma A.3 shows that $g(y^*) < 1$, we are done. Suppose $g^* < 1$, so $g^* = \gamma_1/c$. Plugging in the definition of y^* from equation (5) into equation (A.1) from Proposition A.1 for $g(y^*)$ yields

$$g(y^*) = \frac{\gamma_1}{c + \gamma_3\alpha},$$

where we also use $g(y^*) < 1$. Since γ_1 and γ_3 are strictly positive, it follows that $g(y^*) < \gamma_1/c = g^*$. □

Covid Economics 65, 20 January 2021: 1-46

A.4 Proposition 3

First, since the Lagrangian is linear in v and σ and since individual types are drawn from a continuous distribution F with a pdf f, we can restrict our attention to pure vaccination and interaction policies, i.e. for all (x, y), $v(x, y) \in \{0, 1\}$ and $\sigma(x, y) \in \{0, 1\}$.

Now, for a type (x, y) the planner's vaccination and interaction policy decision is to maximize over the four possible choices for $\sigma(x, y)$ and $v(x, y)$ in $\{0, 1\}^2$. Define $d_{i,j}$ to be the difference between the Lagrangian when $\sigma(x, y) = i$ and $v(x, y) = j$ and the Lagrangian when $\sigma(x, y) = v(x, y) = 0$. We have that $d_{0,0} = 0$, $d_{0,1} = -\gamma_1 dF(x, y)$,

$$d_{1,0} = (xc - \lambda xyb)f(x, y) - \gamma_3 \alpha x f(x, y)$$

and if $\sigma(x, y) = v(x, y) = 1$, the difference is

$$d_{1,1} = xcf(x, y) - \gamma_1 f(x, y).$$

Since $d_{0,1} < 0$, it cannot be optimal for the planner to set $\sigma(x, y) = 0$ and $v(x, y) = 1$. If $y < (c - \gamma_3\alpha)/(\lambda b)$, then $d_{1,0} > 0$ and $d_{1,1} > d_{1,0}$ if and only if $x > \gamma_1/(\lambda yb + \gamma_3\alpha)$. If the opposite inequality for y holds, then $d_{1,0} < 0$ and $d_{1,1} > 0$ if and only if $x > \gamma_1/c$. It follows that if $y < (c - \gamma_3\alpha)/(\lambda b)$, then the individual interacts, $\sigma(x, y) = 1$, regardless of x and the individual is vaccinated, $v(x, y) = 1$, if and only if $x > \gamma_1/(\lambda yb + \gamma_3\alpha)$. If $y > (c - \gamma_3\alpha)/(\lambda b)$, then the individual interacts and is vaccinated, $\sigma(x, y) = v(x, y) = 1$, if $x > \gamma_1/c$. Otherwise, the individual does not interact and is not vaccinated. Therefore, an unvaccinated individual interacts if and only if

$$y > y^* = \min\left\{\frac{c - \gamma_3\alpha}{\lambda b}, 1\right\}$$

and regardless of x. If $y \leqslant y^*$, then an individual is vaccinated if and only if

$$x > g(y) = \min\left\{\frac{\gamma_1}{\lambda by + \gamma_3\alpha}, 1\right\}.$$

If $y > y^*$, then an individual is vaccinated if and only if $x > g(y^*) = \gamma_1/c$.

Finally, the same arguments as in the proofs of Lemmas A.3 and A.4 show that γ_1 and γ_3 are strictly positive. Since the x-threshold function is continuous, the same argument as in the proof of Corollary 4 shows that the optimal vaccination policy is monotone and exhibits an exposure premium everywhere.

A.5 Corollary 1

Given a flat interaction tax, τ, an individual wants to interact if and only if $\lambda by + \tau < c$. If $y^* < 1$, then $\lambda by^* = c - \gamma_3\alpha$. Since $\gamma_3 > 0$, the result follows. If $y^* = 1$, then $\lambda b \leq c - \gamma_3\alpha$. Since $\gamma_3 > 0$, it follows that all individuals want to interact, as in the first best.

Covid Economics 65, 20 January 2021: 1-46

A.6 Lemma 2

The left-hand-side of equation 16, $p/(bA^*)$, is decreasing and continuous in A^*, goes to p/b as A^* goes to 1, and goes to positive infinity as A^* goes to 0. The right-hand-side is increasing and continuous in A^*, and is always weakly positive. It follows that if the right-hand-side is strictly greater than p/b at $A^* = 1$, then there exists an unique equilibrium with $A^* \in (0, 1)$. Otherwise, the left-hand-side is always strictly greater than the right-hand-side for $A^* < 1$ and the unique equilibrium has $A^* = 1$. We can never have $A^* = 0$ because types in an open neighborhood of $(0, 0)$ gain almost nothing from accepting vaccination, but lose $p > 0$.

A.7 Proposition 4

If $A^* = 1$, then the proposition trivially holds with $\bar{y} = 0$. As such, for the remainder of the proof, suppose $A^* \in (0, 1)$.

First, the Lagrange multiplier on the vaccine supply constraint, γ_1, is weakly positive; it may be 0 if on the margin, the planner vaccinates an individual who does not accept. Now, suppose a type (x, y) individual will accept a vaccine if offered. The first derivative of the Lagrangian with respect to $v(x, y)$ is

$$\left[\lambda byx - p - \gamma_1 + \gamma_3 \alpha x\right] f(x, y),$$

where the first two terms are the direct benefit and cost of the vaccine to the individual, the third term is the cost of providing the vaccine, and the fourth term captures the externality of lowering the equilibrium transmission rate. The second derivative with respect to $v(x, y)$ is always 0. If $y \geqslant \hat{y}$, then $\lambda by + \gamma_3\alpha \geqslant \gamma_1 + p > 0$. It follows that the derivative is strictly increasing in x, is strictly negative if $x < (\gamma_1 + p)/(\lambda by + \gamma_3\alpha)$, strictly positive if the opposite strict inequality holds, and equal to 0 if the inequality holds with equality. On the other hand, if $y < \hat{y}$ and $\lambda by + \gamma_3\alpha \geqslant 0$, then the derivative is weakly increasing in x and strictly negative at $x = 1$, so it is strictly negative for all x. Finally, if $y < \hat{y}$ and $\lambda by + \gamma_3\alpha < 0$, then the derivative is strictly decreasing in x and strictly negative at $x = 0$, so it is strictly negative for all x. It follows from the definition of $g(\cdot)$ in the proposition that if an individual with $y < \hat{y}$ is willing to accept a vaccine, then the planner optimally sets $v(x, y) = 0$ and if an individual with $y \geqslant \hat{y}$ is willing to accept a vaccine, then the planner optimally sets $v(x, y) = 1$ if and only if $x \geqslant g(y)$. If an individual is not willing to accept a vaccine, then the planner optimally sets $v(x, y) = 1$ if and only if $x \leqslant tA^*/y$, where $t \in [0, 1]$. If $\gamma_1 > 0$, then $t = 0$. These individuals are vaccinated simply to satisfy the vaccine supply constraint. We can equally suppose that the planner need not satisfy the constraint with equality.

What remains to prove the proposition is to show that there exists a $\bar{y}$ such that $g(y)y \geqslant A^*$ if and only if $y \geqslant \bar{y}$, and that for all $y \geqslant \max\{\hat{y}, \bar{y}\}$, the policy exhibits an exposure premium.

To prove these results, we first rewrite the planner's problem as choosing $A^* \in [0, 1]$, $\lambda \in [0, 1]$, a

CEPR PRESS

continuous and decreasing x-threshold function $g : [0, 1] \to [0, 1]$, and $t \in [0, 1]$ to maximize

$$\mathcal{L} = \mathcal{W} + \gamma_1 \left(\beta - \int_0^1 \int_{\max\{g(y), A^*/y\}}^1 f(x, y) dx dy - \int_0^1 \int_0^{tA^*/y} f(x, y) dx dy \right) + \gamma_2 \left(A^* - \min \left\{ \frac{p}{\lambda b}, 1 \right\} \right)$$
$$+ \gamma_3 \left(\lambda - \alpha \int_0^1 \int_0^{\max\{g(y), A^*/y\}} x f(x, y) dx dy \right),$$

where welfare is

$$\mathcal{W} = -\lambda b \int_0^1 \int_0^{\max\{g(y), A^*/y\}} y x f(x, y) dx dy - \int_0^1 \int_{\max\{g(y), A^*/y\}}^1 p f(x, y) dx dy.$$

Before we can proceed, we first have to show that the Lagrange multiplier on the vaccine hesitancy constraint, γ_2, is weakly positive and on the equilibrium transmission rate constraint, γ_3, is strictly positive. To that end, observe that the derivative of welfare with respect to A^* is 0 since A^* is optimally chosen to equate $\lambda b A^*$ with p. We then have that

$$\frac{\partial \mathcal{L}}{\partial A^*} = \gamma_1 \int_{y: A^*/y > g(y)} \frac{1}{y} f(A^*/y, y) dy + \gamma_2 - \gamma_3 \alpha \int_{y: A^*/y > g(y)} \frac{A^*}{y^2} f(A^*/y, y) dy.$$

If the set of y such that $A^*/y > g(y)$ has measure 0, then the First Order Condition for A^* implies that $\gamma_2 = 0$. If that set has strictly positive measure, then for any y in that set, we can increase $g(y)$ to A^*/y without affecting any other variables, including welfare, at which point the derivative of the Lagrangian with respect to $g(y)$ is

$$\frac{\partial \mathcal{L}}{\partial g(y)} = \gamma_1 f(A^*/y, y) - \gamma_3 \alpha \frac{A^*}{y} f(A^*/y, y) dy,$$

where we use that the derivative of welfare with respect to $g(y)$, like the derivative of welfare with respect to A^*, is 0. For the policy to be optimal, it must be that this derivative is weakly negative. Plugging into the First Order Condition for A^*, which sets the derivative of the Lagrangian with respect to A^* equal to 0, implies that $\gamma_2 \geqslant 0$. In any case, we have that $\gamma_2 \geqslant 0$.

If $\gamma_2 > 0$, then since $\gamma_1 \geqslant 0$, the First Order Condition for A^* implies that $\gamma_3 > 0$. For $\gamma_2 = 0$, consider the derivative of the Lagrangian with respect to λ:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -b \int_0^1 \int_0^{\max\{g(y), A^*/y\}} y x f(x, y) dx dy + \gamma_2 \frac{p}{\lambda^2 b} + \gamma_3.$$

The First Order Condition sets this derivative equal to 0, which implies along with $\gamma_2 = 0$ that $\gamma_3 > 0$. As such, in any case, we have that $\gamma_3 > 0$.

Now, to see the remaining results – that there exists the desired $\bar{y}$ such that $g(y)y \geqslant A^*$ if and only if $y \geqslant \bar{y}$ and that the optimal policy exhibits an exposure premium for all $y \geqslant \max\{\hat{y}, \bar{y}\}$ – consider the

Covid Economics 65, 20 January 2021: 1-46

derivative of $g(y)y$ with respect to y:

$$\frac{\partial g(y)y}{y} = \frac{(\gamma_1 + p)\gamma_3\alpha}{(\lambda by + \gamma_3\alpha)^2}.$$

The derivative is strictly positive since $\gamma_1 \geqslant 0$ and $\gamma_3 > 0$. Both results follow. Moreover, if $\bar{y} > 0$, then it is given by $g(\bar{y})\bar{y} = A^*$. Thus, for all $y \in [\hat{y}, \bar{y})$, the policy is an xy-policy with iso-risk threshold A^*.

A.8 Lemma 3

Define $\tilde{\gamma}_1 \triangleq \gamma_1/c$, $\tilde{\gamma}_3 \triangleq \gamma_3/b$, and $\mu \triangleq \lambda/\alpha$. The system of First Order Conditions for the Problem 1 version of the second-best problem (described in Subsection A.3 and in which we assume $y^* < 1$) can be written as the following system of six equations:

$$\beta = \int_0^{y^*}\int_{g(y)}^1 f(x,y)dxdy + \int_{y^*}^1\int_{g^*}^1 f(x,y)dxdy, \tag{A.5}$$

$$\mu = \int_0^{y^*}\int_0^{g(y)} xf(x,y)dxdy, \tag{A.6}$$

$$y^* = \frac{c}{\alpha b}\frac{1}{\mu}, \tag{A.7}$$

$$g(y) = \min\left\{\frac{c}{\alpha b}\frac{\tilde{\gamma}_1}{\mu y + \tilde{\gamma}_3}, 1\right\}, \tag{A.8}$$

$$g^* = \min\{\tilde{\gamma}_1, 1\}, \tag{A.9}$$

and

$$\tilde{\gamma}_3 = \frac{\int_{g(y^*)}^{g^*} xf(x,y^*)dx - \tilde{\gamma}_1\int_{g(y^*)}^{g^*} f(x,y^*)dx + \left(\frac{\alpha b}{c}\right)^2\mu^2\int_0^{y^*}\int_0^{g(y)} yxf(x,y)dxdy}{\left(\frac{\alpha b}{c}\right)^2\mu^2 + \frac{\alpha b}{c}\int_0^{g(y^*)} xf(x,y^*)dx}, \tag{A.10}$$

which we solve for μ, y^*, $g(\cdot)$, g^*, $\tilde{\gamma}_1$, and $\tilde{\gamma}_3$. The final equation comes from setting

$$\frac{\partial\mathcal{L}}{\partial y^*} - \frac{\lambda^2 b}{c}\frac{\partial\mathcal{L}}{\partial\lambda} = 0;$$

we can then use the First Order Condition for either λ or y^* to back out γ_2. Finally, note that

$$\mathcal{W} = -c\left(\frac{\alpha b}{c}\mu\int_0^{y^*}\int_0^{g(y)} yxf(x,y)dxdy + \int_{y^*}^1\int_0^{g^*} xf(x,y)dxdy\right),$$

so welfare scales with the interaction avoidance cost, c, but only $\alpha b/c$ affects welfare comparisons across policy choices.

For Problem 2, a similar argument applies except we no longer need equation (A.7) for y^* or equation

Covid Economics 65, 20 January 2021: 1-46

(A.9) for g^*, and equation (A.10) for $\tilde{\gamma}_3$ becomes

$$\tilde{\gamma}_3 = \int_0^1 \int_0^{g(y)} yxf(x,y)dxdy.$$

For the first-best, a similar argument applies except we no longer need equation (A.9) for g^*, we replace equation (A.5) with

$$\beta = \int_0^{y^*} \int_{g(y)}^1 f(x,y)dxdy + \int_{y^*}^1 \int_{g(y^*)}^1 f(x,y)dxdy, \tag{A.11}$$

we replace equation (A.7) with

$$y^* = \min\left\{ \frac{\frac{c}{\alpha b} - \tilde{\gamma}_3}{\mu}, 1 \right\}, \tag{A.12}$$

and we replace equation (A.10) with

$$\tilde{\gamma}_3 = \int_0^{y^*} \int_0^{g(y)} yxf(x,y)dxdy. \tag{A.13}$$

B Computational Methods

To compute the first- and second-best optimal policies, we compute solutions to the systems of equations described in Subsection A.8. In each case, we show that the system of equations minus one has an unique solution given a guess for $\tilde{\gamma}_3/\mu$. We then iterate over $\tilde{\gamma}_3/\mu$ until the final equation is satisfied. To solve the second-best problem, we conduct this procedure for Problem 1 (assuming $y^* < 1$) and Problem 2 (assuming $y^* = 1$) and choose the solution that achieves the highest welfare. To solve the first-best problem, we simply choose the solution that achieves the highest welfare. In all our numerical experiments, the two sides of the final equation are continuous and monotone in $\tilde{\gamma}_3/\mu$ and in opposite directions, so there is clearly an unique solution to the system of First Order Conditions.

B.1 Second-best Policy

Lemma B.1. *Taking as given $\tilde{\gamma}_3/\mu$, the system of the first five equations in Lemma 3, equations* (A.5)-(A.9)*, has an unique solution for the remaining five unknowns, μ, y^*, $\tilde{\gamma}_1$, $g(\cdot)$, and g^*.*

Proof. First, we show that taking as given $\tilde{\gamma}_1/\mu$ and $\tilde{\gamma}_3/\mu$, the system of four equations, (A.6) - (A.9), i.e. the five equations minus the vaccine supply constraint, has an unique solution for the remaining four unknowns, μ, y^*, $g(y)$, and g^*. Take as given $\tilde{\gamma}_1/\mu$ and $\tilde{\gamma}_3/\mu$. For each value of μ, we know y^* from equation (A.7), $g(y)$ from (A.8), and g^* from (A.9). Hence, it is sufficient to show that there cannot be two solutions to the system of four equations with different values for μ. Suppose there are two solutions (whose variables are indexed by subscripts 1 and 2) with $\mu_1 > \mu_2$. It follows that $y_1^* < y_2^*$, and $g_1 = g_2$. It follows that the right-hand-side of equation (A.6) for the equilibrium value of μ is strictly smaller in

Covid Economics 65, 20 January 2021: 1-46

the first solution than in the second. This contradicts $\mu_1 > \mu_2$. As such, taking as given $\tilde{\gamma}_1/\mu$ and $\tilde{\gamma}_3/\mu$, the system of four equations, (A.6) - (A.9), has an unique solution for the four unknowns, μ, y^*, $g(y)$, and g^*.

Moreover, suppose we choose two different values $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$. Since $\tilde{\gamma}_3/\mu$ is fixed, we have that $g_1 > g_2$. It must be that $\mu_1 > \mu_2$. If $\mu_1 \leq \mu_2$, then $y_1^* \geq y_2^*$, which, along with $g_1 > g_2$, implies an increase in the right-hand-side of equation (A.6) for μ, contradicting $\mu_1 \leq \mu_2$.

Now, take as given $\tilde{\gamma}_3/\mu$ and consider the system of five equations, (A.5)-(A.9), with five unknowns, $\tilde{\gamma}_1/\mu$, μ, y^*, $g(y)$, and g^*. Since the list four equations uniquely define μ, y^*, $g(y)$, and g^* given $\tilde{\gamma}_1/\mu$, it is sufficient to show that there cannot be two solutions with different values for $\tilde{\gamma}_1/\mu$. Suppose there are with $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$. We know that $g_1 > g_2$ and $\mu_1 > \mu_2$, which implies that $y_1^* < y_2^*$ and, along with $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$, implies that $g_1^* > g_2^*$. The change from μ_1 to μ_2 has no direct effect on the right-hand-side of the vaccine supply constraint, equation (A.5), but the other three changes each imply that the right-hand-side is strictly lower in the first case than in the second (to see the effect of the change in y^*, recall that $g(y^*) < g^*$). Since the left-hand-side, β, is the same in both cases, this cannot be. It follows that given $\tilde{\gamma}_3/\mu$, the system of five equations, (A.5)-(A.9), has an unique solution for the five unknowns, $\tilde{\gamma}_1/\mu$, μ, y^*, $g(y)$, and g^*.

Furthermore, suppose there are two solutions to the full system of six equations, (A.5)-(A.10), with $(\tilde{\gamma}_3/\mu)_1 > (\tilde{\gamma}_3/\mu)_2$. It must be that $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$. Suppose not. Then $g_1 < g_2$ and it follows that $\mu_1 < \mu_2$. Otherwise, if $\mu_1 \geq \mu_2$, then $y_1^* \leq y_2^*$, which, along with $g_1 < g_2$, implies that the right-hand-side of the equilibrium transmission rate constraint, (A.6), is strictly smaller in the first case than in the second, which contradicts $\mu_1 \geq \mu_2$. As such, $\mu_1 < \mu_2$, which implies that $y_1^* > y_2^*$ and, along with $(\tilde{\gamma}_1/\mu)_1 \leq (\tilde{\gamma}_1/\mu)_2$, implies that $g_1^* < g_2^*$. It follows that the right-hand-side of the vaccine supply constraint, (A.5), is strictly larger in the first case than in the second, which cannot be since β is the same across the two cases. As such, we have that $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$.

Finally, we show that $\tilde{\gamma}_3/\mu \in [0, 1]$. Since we showed in Lemma A.4 that $\gamma_3 > 0$, it follows that $\tilde{\gamma}_3/\mu > 0$. Moreover, consider the sixth equation in our system, (A.10) for the combined First Order Conditions for λ and y^*. Since $\tilde{\gamma}_1 \geq g^*$, we can see that

$$\int_{g(y^*)}^{g^*} x f(x, y^*) dx < \tilde{\gamma}_1 \int_{g(y^*)}^{g^*} f(x, y^*) dx.$$

Moreover, $\int_0^{g(y^*)} x f(x, y^*) dx > 0$. It then follows from plugging in the equilibrium value of μ that

$$\frac{\tilde{\gamma}_3}{\mu} < \frac{\int_0^{y^*} \int_0^{g(y)} y x f(x, y) dx dy}{\int_0^{y^*} \int_0^{g(y)} x f(x, y) dx dy},$$

where the right-hand-side is the average value of y among those in the interaction pool. Since all individuals have $y \leq 1$, it follows that this average is less than 1. □

B.2 First-best Policy

Lemma B.2. *Take as given $\tilde{\gamma}_3/\mu$. There is an unique solution to the system of four equations (excluding equation* (A.13) *for $\tilde{\gamma}_3$),* (A.6)*,* (A.8)*,* (A.11)*, and* (A.12)*, for the remaining four unknowns, $\tilde{\gamma}_1$, μ, y^*, and $g(\cdot)$.*

Proof. We first show that taking as given $\tilde{\gamma}_3/\mu$ and $\tilde{\gamma}_1/\mu$, there is an unique solution to the system of three equations (excluding the vaccine supply constraint), (A.6), (A.8), and (A.12), for the remaining three unknowns, μ, y^*, and $g(\cdot)$. Note that g is fully determined by $\tilde{\gamma}_1/\mu$ and $\tilde{\gamma}_3/\mu$. It is then sufficient to note that equation (A.6) implies that μ is strictly increasing in y^* and equation (A.12) implies that y^* is decreasing in μ.

To complete the proof of the lemma, it is sufficient to show that there cannot be two solutions to the system of four equations, (A.6), (A.8), (A.11), and (A.12), with different values for $\tilde{\gamma}_1/\mu$. Suppose there are two such solutions, indexed by subscripts 1 and 2, with $(\tilde{\gamma}_1/\mu)_1 > (\tilde{\gamma}_1/\mu)_2$. It follows that $g_1 > g_2$. If $\mu_1 \leq \mu_2$, then equation (A.12) shows that $y_1^* \geq y_2^*$. But then equation (A.6) implies that $\mu_1 > \mu_2$, a contradiction. It follows that $\mu_1 > \mu_2$ and $y_1^* < y_2^*$. Now, the right-hand-side of equation (A.11) is strictly decreasing in g and strictly increasing in y^*. To see the second claim, observe that the derivative of the right-hand-side with respect to y^* is $-\int_{y^*}^{1} g'(y^*) f(g(y^*), y) dy$, which is strictly positive since the derivative of g with respect to y is always strictly negative. It follows that the right-hand-side is strictly smaller in the first solution than in the second, which cannot be the case since the left-hand-side, β, is the same in both cases. □

Covid Economics 65, 20 January 2021: 1-46

Covid Economics Issue 65, 20 January 2021

Schools opening and Covid-19 diffusion: Evidence from geolocalized microdata[1]

Emanuele Amodio,[2] Michele Battisti,[3] Andros Kourtellos,[4] Giuseppe Maggio[5] and Carmelo Massimo Maida[6]

Date submitted: 7 January 2021; Date accepted: 12 January 2021

Are schools triggering the diusion of the Covid-19? This question is at the core of an extensive debate about the social and long-run costs of stopping the economic activity and human capital accumulation from reducing the contagion. In principle, many confounding factors, such as climate, health system treatment, and other forms of restrictions, may impede disentangling the link between schooling and Covid-19 cases when focusing on a country or regional-level data. This work sheds light on the potential impact of school opening on the upsurge of contagion by combining a weekly panel of geocoded Covid-19 cases in Sicilian census areas with a unique set of school data. The identication of the effect takes advantage of both a spatial and time-variation in school opening, deriving by the flexibility in opening dates determined by a Regional Decree, and by the occurrence of a national referendum, which pulled a set of poll-station schools towards opening on the 24th of September. The analysis finds that census areas where schools opened before observed a signicant and positive increase of Covid-19 cases between 1.5-2.9%. This result is consistent across several specications, including accounting for several schools opening determinants, such as the number of temporary teachers, Covid-19 cases in August, and pupils with special needs. Besides,

1 We are extremely grateful to Francesco Vitale for his support on this research. We wish also to thank Massimiliano Sacco and Francesco Raul Ciaccio for their assistance with the school data collection. We also thank the support of the Italian COVID-19 integrated surveillance platform, coordinated by the Department of Infectious Diseases of the Istituto Superiore di Sanita, for the access to the Covid-19 data. All errors remain ours.

2 Department of Health, Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo.
3 Department of Law, University of Palermo.
4 Department of Economics, University of Cyprus.
5 School of Global Studies, University of Sussex
6 kDepartment of Health, Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo.

Covid Economics 65, 20 January 2021: 47-77

school opening increases non-linearly the number of instances in zones observing already some Covid-19 cases. Using the estimated coecients, this paper also presents a prediction of Covid-19 cases with dierent school opening scenarios. Finally, an exploration of the heterogeneity at school and demographic levels, including class size and age, poses the basis for calibrated policies to control dierential reopening.

Covid Economics 65, 20 January 2021: 47-77

Covid Economics 65, 20 January 2021: 47-77

1 Introduction

Since last October, a harsh second wave of the Covid-19 pandemic has been hitting the economic, health, and educational systems of many countries worldwide. A major challenge of policymakers is to understand the role of these sectors on the diffusion process of Covid-19. With an estimated drop in GDP spanning between 14.7 and 32.9 percent in the EU and the US (e.g., Chudik et al. (2020)), identifying and calibrating a set of policies that still avoid large scale shutdowns while minimizing the levels of contagion is vital.

Among the set of early actions undertaken by policymakers was to implement school closures since children are often considered as a major source of contagion. Figure 1 shows the time pattern of smoothed new Covid-19 cases, weighted by millions of inhabitants for 6 OECD countries, and the first day of school opening registered in each country.[1] The evidence appears to be mixed: while the cases increased after mid-September occurred in all the countries, the first day of school varies sensibly, with countries that started schools earlier as Germany and others where the increase in cases and school opening was almost concomitant, such as Italy and Spain.

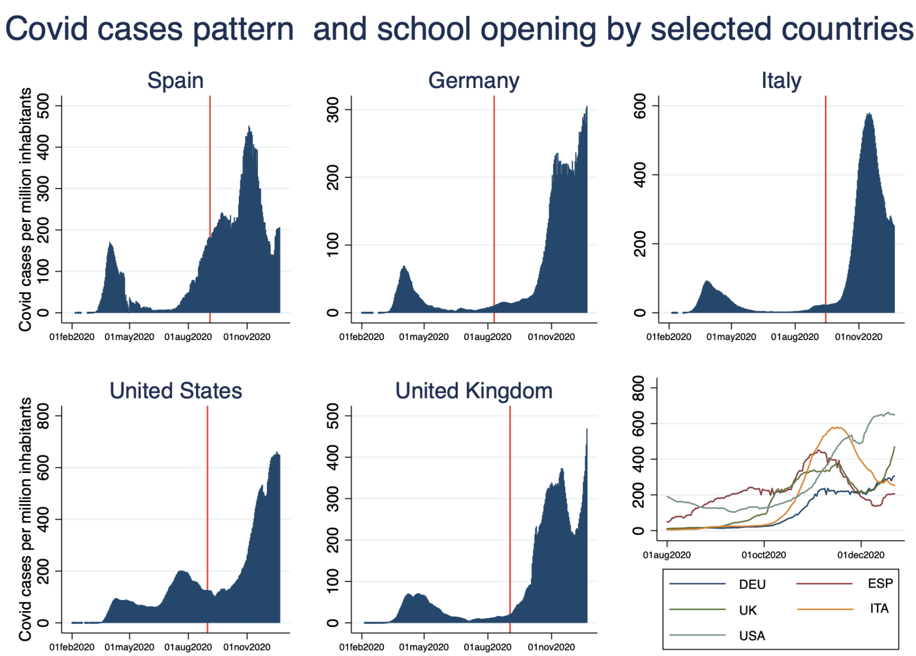


Figure 1: Covid-19 cases and school opening by countries

However, the efficacy of school closures has been the subject of intense debate in policy and academic circles. The difficulty to reach a consensus about the role of school on Covid-19 diffusion is multifold. First, when the pandemic hit the world in the early months of 2020, the knowledge of health professionals about the new disease was limited and governments were unprepared to deal with a pandemic that constitutes an

[1] The data were obtained from https://ourworldindata.org/coronavirus and the first day of school is defined as the first day of school of at least a region in the country.

unprecedented shock for the world at least in the modern era. As a result, the first phase of intervention rightly focused on reducing the health cost of the diffusion by implementing strict lockdowns, including shutting down the face-to-face classes without a precise estimate of the impact of these policies on Covid-19 diffusion. Second, the symptomatology of Covid-19 itself may hinder a proper identification of the virus among children, as these are more often asymptomatic and thus less likely to get tested. Third, drawing conclusions from extant studies on similar diseases may not help to find a solution. One reason is that while the literature on the role of schools on the diffusion of influenza is large, the wide availability of influenza vaccines may impede direct comparisons with the case of Covid-19 for which vaccines have just started to be available to special groups. Another reason for the inability of influenza studies to provide useful information on SARS type viruses like Covid-19 and common influenza is due to their differences in incubation and serial periods that determine the speed of transmission of the viruses.[2]

This paper investigates the role of schooling in the diffusion of Covid-19 by exploiting exogenous variation, varying over both time and space, in the school opening from a natural experiment using Italian micro data from Sicily. The idea is to exploit spatial variation in the opening of schools due to a sudden change in regulation from the Regional Government and the occurrence of a national referendum, held on September 20th-21st. Since some schools were used as polling stations, the school opening date varied from few days before the planned official opening date to some date after the referendum generating granular space-time heterogeneity which allows us to identify the differential effect of school opening on the diffusion of Covid-19.

Our paper contributes to the current literature in several ways. First, our study is instrumental to the literature on present and future social costs of closing schools. Some recent works show as these costs may be huge from the side of human capital losses. A recent macro study at country levels from Psacharopoulos et al. (2020) estimates a loss of 8% in future earnings that match a similar loss in aggregate human capital. Other studies with micro data on school data achievement (but without Covid-19 data) show similar evidence. Engzell et al. (2020) find a loss in learning equal to about 3%, measured through the final tests conducted in Dutch primary schools right after the first wave and lockdown of Covid-19. Agostinelli et al. (2020) examine the effects of school closures during the Covid-19 pandemic on children's education by considering the interaction of schools, peers, and parents. Using the Add Health data, they find that school closures have an asymmetric, large and persistent effect on educational outcomes, leading to higher educational inequality.

Second, more directly, this work deals with the effects of school opening as trigger of Covid-19 increases. A growing strand of literature has started to address this relationship using various approaches and levels of time and space definition. One approach is based on direct comparisons between cases within the school system with respect to general population. On this line of work, the survey of Lewis (2020) and the work of Oster (2020) suggest low rates of contagion in US schools, contingent to students participation to the survey. Similarly,

[2] In terms of the effect of school closure on other diseases, works focusing on influenza show that these imply a reduction in the transmission of the virus ranging from 4.2% for influenza B virus epidemic in Hong Kong in 2018 (Ali et al. (2018)) to 8.2% for Pandemic Flu H1N1-2009 in Oita (Kawano and Kakehashi (2015)). A review of historical evidence by Cauchemez et al. (2009)), based on school holidays in France and the 1918 experience of US cities, suggest a 15% reduction, but with the potential of larger reductions if the children are isolated and policies well implemented.

Covid Economics 65, 20 January 2021: 47-77

Covid Economics 65, 20 January 2021: 47-77

Buonsenso et al. (2020) study total infections in Italian schools and find that less than 2% of schools show infections at the date of October 5th. Sebastiani and Palú (2020) find an upward trend of Italian cases few weeks after the school openings. Another approach exploits the differences in the timing of school opening. For example, Isphording et al. (2020) find no evidence on schools as triggers of Covid-19 upsurge when looking at the time discontinuity in schools opening among German landers. For the Italian case, Lattanzio (2020) exploits the differences in school reopening among Italian regions and finds a positive relationship between earlier opening and regional variation in total Covid-19 cases. In the same line of work, Gandini et al. (2020) focus on the growth of Covid-19 incidence across all age groups after school opening in Veneto, finding that this was lower in school age individuals, while maximal in 20-29 and 45-49 years old individuals.

The aforementioned approaches are likely to give rise to results that are subject to several biases. One potential source of bias are the students themselves, as these are more likely to be asymptomatic than the the older population. Therefore, one may argue that the number of cases tested and detected are much lower than the ones in the remaining population.[3] What is more, students may act as triggers of contagion within their households or social network, and this may be likely to reduce the difference between treated (students) and controls (population) while increasing the overall level of contagion.

Another concern is that analyses based on aggregated data at regional level are unable to account for idiosyncratic elements, such as the heterogeneity in testing or healthcare management, or other related factors. More precisely, the starting date at regional level is based on a political decision, which may depend on the level of Covid-19 at the time of the decision and the number of cases in the relevant territory or among the employees. Not accounting for these factors would result in identifying spurious relationships driven by other hidden effects such as bureaucratic efficiency. Furthermore, the decision of opening schools may be delayed by school-managers according to a set of local regulations, or when schools are seats of polling stations during elections in the early stage of a school year. Age groups analyses, in addition, are very likely to suffer of the young population bias, due to the absence of a true counterfactual group. Unfortunately, some of these issues remain unsolved when comparing different policies at national level, which are often unable to capture the local level volatility of their implementation.[4] In contrast, our study is not subject to the above limitations because it relies on disaggregate micro data and exploits a random variation in the opening of the schools due to the national referendum.

In particular, the present work contributes to the recent literature on Covid-19 and the educational sector by shedding light about the role of schools opening on the recent upsurge of Covid-19 cases at local level. The analysis employs a dynamic panel model based on a unique panel data of Covid-19 cases geocoded at census area level for the region Sicily, consisting in blocks of about 150 inhabitants. These data are merged with granular geocoded data on school opening date, collected from the online official documents (*Circolare*) and

[3]A summary review includes for instance Li et al. (2020), Mehta et al. (2020), and Maltezou et al. (2020).

[4]For example, Haug et al. (2020) develop a comprehensive study on the effect of non medical interventions on the effect of the closure of educational institutions in reducing Rt ratio. Hsiang et al. (2020) estimate an average reduction of 11% on infection growth rate by school closures in Italy during the first wave of March 2020.

communication to parents of each of the 4,223 public schools in Sicily. The identification is based on space and time variation of school opening, while controlling for unobserved time invariant heterogeneity, time dummies, and the lagged level of Covid-19 cases in the same census area. Spatially, this work combines Covid-19 data at census area level with date of opening of the schools within 1 km, as this is the average travel distance of students up to the secondary schools.

The results suggest that census areas anticipating school opening observed a strongly significant differential positive effect on Covid-19 cases, ranging between 1.5-3% in the specifications for the cases of the whole population. The effect is lower for population under 19 years, supporting the asymptomatic hypothesis of the young cohort of population, while it is higher for the remaining cohorts. Also, the effect appears larger for areas with schools holding larger class sizes, and for less densely populated census areas, where schools are likely to represent the major source of social interactions. Finally, using the estimated coefficients, the final part of the paper presents a set of alternative scenarios offering the magnitude of decrease in cases if the school would have not reopened, or if contacts within the schools would have contained more.

The work is organized as follows. Section 2 provides background information about the Covid-19 diffusion in Sicily and Section 3 describes the data. Section 4 specifies the econometric model, Section 5 presents the results, and Section 6 concludes.

2 Background: Covid-19 diffusion in Sicily

Sicily was just marginally affected by the first Covid-19 wave of February-May 2019, with only 2,735 positive cases registered between February and May, a level much lower than the rest of Italy. This analysis relies on data from Istituto Superiore di Sanità (ISS) collected and provided, on a daily basis, by the Sicilian Region. These contain anonymized information on daily new cases, age sex of the individual, and a dummy whether they are linked to the educational sector.[5] It is possible to notice that from the end of September onward, the pattern of diffusion increased substantially, becoming most similar to the other most affected regions in Italy. The cases reach a peak of 1,871 new cases on November 9th, equal to about 68 percent of all the Sicilian cases during the first wave of Covid-19 (see Figure 2). The increase observed in late September may have many explanations. On one side, national public opinion has pointed to the decrease in restrictions and enforcement of controls that occurred during summer, implemented to save the tourist season, which in Sicily represents one of the major industry. On the other side, the timing of the increase has been aligned with school opening, pushing some observers in pointing towards the role of the school. Finally, seasonality may have determined an increase in Covid-19 cases, which as other coronaviruses proliferate in colder environments as summarized in Carlson et al. (2020).

According to the data, the total cumulative cases December 14th are more than 70,000, but some consist of individuals with residence outside Sicily, such as tourists and commuters, who are excluded from geocoding.

[5]These include both students and schools employees, such as teachers and other staff.

Covid Economics 65, 20 January 2021: 47-77

CEPR PRESS

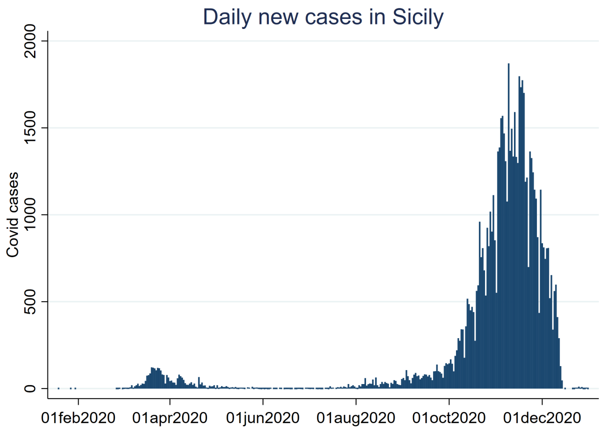


Figure 2: New daily Covid-19 cases in Sicily

The total cases of residents on the island are about 69,107 cases, a number that considered per capita is in line with the rest of the Southern regions and about half of the average in the Northern ones. Figure 3 reports the cumulative cases from September 1st onwards and suggest that, out of all the resident cases, 59,899 are people aged 20 or above, while the remaining 9,208 are 19 or younger. School-related cases are 1,391, all of them from September 1st onward, since this information has been collected only starting from this date. Despite the differences in the cumulative trend, the cumulative growth of cases remains quite similar across the groups, as shown by Figure A1 in Appendix.

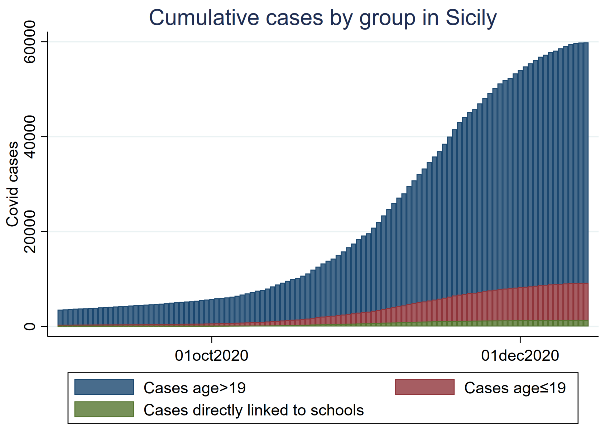


Figure 3: Cumulative cases by group from September onwards.

In terms of diffusion of the virus, at the regional level, the weekly Rt reproduction index slightly increases by 0.4 points from late August until early October, passing from 0.82 on August 24th-30th, up to 1.22 on September 28th-October 4th (Istituto Superiore di Sanità (ISS)).[6] Despite the growth of cases shows an

[6]Table A3 in Appendix show the Rt series including the confidence intervals, which almost always overlap for the period under

Covid Economics 65, 20 January 2021: 47-77

upward trend in concomitance with the opening of the school, this evidence needs to be taken with caution because of all the aforementioned challenges in identification. An additional confounding factor may derive from a dramatic change in testing in September, which may have determined an increase in the positive cases spotted in concomitance with school opening. As Figure 4 shows, this does not seem to hold for the case of Sicily: while the growth rate of cases has increased sharply from September onwards, the growth rate of tests has remained constant, pointing to an increase in the share of positive found per number of tests.[7]

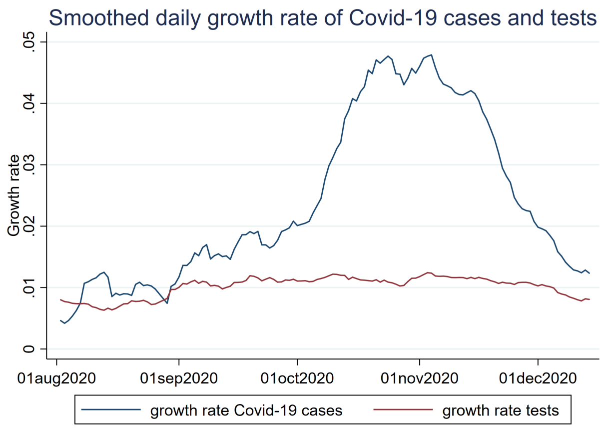


Figure 4: Growth rate of Covid-19 cases and testing.

This evidence indicates that keeping everything else constant, the number of cases increased after September. Also, this suggests that any result from the present analysis is robust to sudden changes in the numbers of tests.

3 Data

The analysis involves a unique panel data at census area level for August 1st-December 14th, obtained by merging daily Covid-19 cases in Sicily with information on public school opening for the academic year 2020-21. In the benchmark set of specifications we use a weekly panel with 33,666 observations for 20 weeks, becoming 18 due the dynamic framework of our estimations. The standard setup contains 597,619 observations. The unit of analysis is the census area, which is the smallest granular administrative area, roughly corresponding to blocks.[8] According to the 2011 National Census conducted by the Istituto Nazionale di Statistica (ISTAT), each unit contains on average 152 and a maximum of 3,036 residents. These units have a median area of 0.13 km^2 for inhabited zones, for an average population density of about 714.71 inhabitants/km^2.[9] Table A2 in

study.

[7]The growth rate have been smoothed using a 3 day moving average to account for daily differences in testing.

[8]Sicily comprises 36,681 census areas, but ISTAT reports inhabitant only for 33,366 of these.

[9]The last official census dates back to 2011. While some estimates are available at regional and municipality level, the census-area population estimates are not available, also due to the high degree of migration that occurred in 10 years. For this reason, the analysis uses the census areas only to define the unit of observation and does not take the cases as population shares. Aware about

Covid Economics 65, 20 January 2021: 47-77

Appendix reports the summary statistics of the variables we use in the estimations of section 4.

3.1 Covid data

Our dependent variable is the geolocalized total weekly Covid-19 cases at the census area obtained from Istituto Superiore di Sanità (ISS), which is the office monitoring the Covid-19 pandemic.[10] Our choice to aggregate the data at the weekly frequency follows the standard practice in the literature to account for the serial time of infection that is 7 days (e.g., Cereda et al. (2020)).

In particular, selecting seven days accounts for the incubation time, which takes about five days, and the additional time to conduct the testing and receive the results. In addition, to account for the dynamic evolution of contagion at the census area level, the empirical specification will include the first and second weekly lag of the Covid-19 cases in the same census area. Both the dependent and the lagged variables are measured in natural logarithms, after adding one to each observation.[11]

Using these population data, Figure 5 reports the quintile map of cumulative Covid-19 cases over population at municipality level for August 1st-December 14th. While using the municipality level helps to evidence eventual spatial clusters within Sicily, in the analytical part the unit of analysis remains the census areas, which are in a magnitude order of 1:100 with respect to municipalities. The most densely populated municipalities, including Palermo in the north-west, Catania on the east, and Siracusa in the south-east, observe the highest rate of Covid-19 cases per 1,000 inhabitants, spanning between 15.72 and 40.33 for the period under consideration.[12] Figure A3 in the Appendix offers a disaggregated picture of the unit of analysis and the dependent variable by visualizing Covid-19 cases at census-area level for Palermo, measured around the peak of cases.

Weekly indicators at census areas are then merged to dates on school opening, the precise school location, which is obtained by extracting the latitude and longitude of the school official address.

3.2 Construction of the indicator on public school opening

Our key explanatory variable is the indicator on public school opening. We construct this variable using information on the particular initial day of the 2020-21 school year for 4,223 public schools in Sicily listed both in the official school list of the Ministry of Education and in the one of the Regional Department of Education. The analysis gathered this information from all the schools' public documents, including official communication to parents and internal directives ("Circolari"). From the end of August onwards, school managers are obliged to communicate to the public the precise date of the school reopening in September.[13]

this limitations, for the heterogeneity analysis on population in section 4.2, the missing information on the census-area population of the 2011 National Census is integrated with the spatial re-elaborated data on population estimates from Worldpop for 2020 (Tatem (2017)).

[10] The Covid-19 database includes individual-level anonymized data of daily new positive cases, including age, sex, census area of residence, and whether they are directly linked to the education sector. The total number of geolocalized cases is 69,107, as for an average of 0.09 cases for census areas between August 1st- December 14th.

[11] For robustness purposes, we also considered a transformation, using the Inverse Hyperbolic Sine Transformation developed by Bellemare and Wichman (2020). Our main results remain unaltered.

[12] Figure A2 in Appendix report the cases for February 26th-December 14th.

[13] Collecting the data has involved a thorough screen of all the documentation uploaded on each school website. This information has been integrated with direct calls to the schools from a team of research assistants.

Covid Economics 65, 20 January 2021: 47-77

Covid Economics 65, 20 January 2021: 47-77

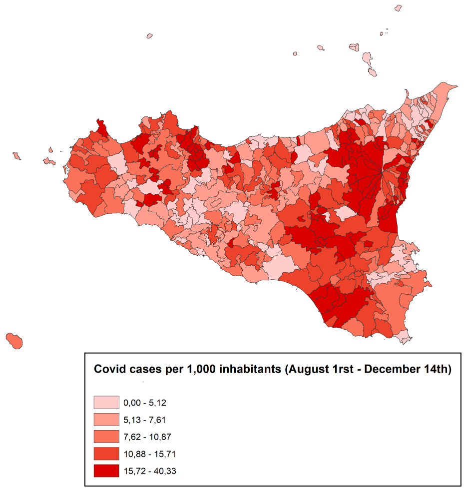


Figure 5: Cumulative cases at December 14th at commune level

The 2020-21 school opening in Sicily shows a huge degree of unforeseen variability due to a set of unexpected events linked to Covid-19 diffusion combined with a national referendum. With a first decree dated August 20th, the Regional Government of Sicily has determined that private schools could have opened on September 1st, while public schools should have started the school year on September 14th, with the only exception of public schools that were polling stations for the national referendum of September 20th-21st, which had the option to start from September 24th. Only few days before the official opening, on August 31st, Sicily has allowed all the schools to set up their schedule with a second decree. This decree included the possibility of opening even after September 24th (see Figure 6 and 7). Some school managers have delayed the opening time to assess the final settlements for the emergency situation. Other school managers, however, stuck to the original plan due to the short notice of the decree and the minimum number of school days to be conducted in a school year, which remained constant.

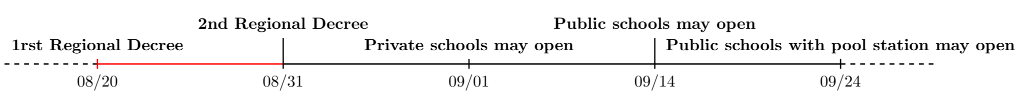


Figure 6: Schools starting dates in Sicily

This change in regulation has determined a large variability in the school year's starting date, with schools opening from September 1st until October 9th. Figures 7 show the distribution of observed school opening weighted by students across all the possible range, suggesting that school managers' decisions were bimodal with the modes around the original date (September 14th) and the new "poll station" date (September 24th).[14]

From an analytical perspective, this set of unexpected events has inadvertently generated a reasonable time

[14] We obtain similar distributions when we plot raw dates or when weighting for the number of students and censoring ten days of right and left tails.

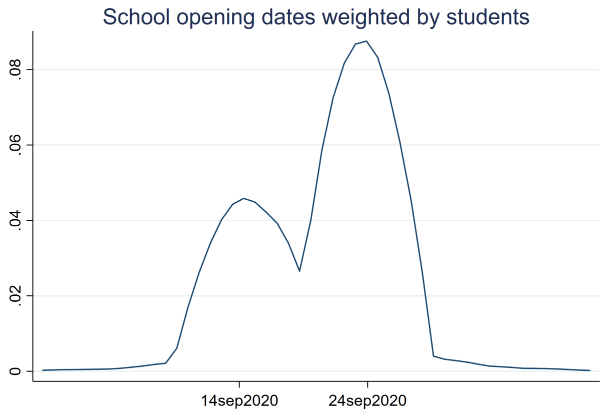


Figure 7: Date of first day in class

Covid Economics 65, 20 January 2021: 47-77

discontinuity adapt to measure Covid-19 diffusion through school, larger both than the median incubation time of 4-5 days (McAloon et al. (2020)) and of the serial interval time of 7.5 days (Cereda et al. (2020)). The latter is defined as the time between a primary case-patient with symptom onset and a secondary case-patient with symptom onset. The time discontinuity persists when focusing on the different levels of schooling (see Figure 8) but changes depending on the level itself. Infancy, primary and middle schools are more likely to be selected as electoral poll stations, thus following the general path of opening, with a large share of school opening around September 24th and a lower share around September 14th. Secondary schools, instead, are less likely to be poll stations, and were more likely to start the school year as originally planned, on September 14th. For example, in Palermo's regional capital, out of 610 polling stations only seven are located in three secondary schools and the other seven in hospitals. All the remaining electoral poll stations are infancy, primary, middle schools and general institutes involving all three types of schools. [15] However, official statistics underscore that only about a third of the public schools are seats of polling stations, suggesting that many school managers decided to start the school year on or after September 24th, contrary to what was originally planned by the first Regional Decree. Accounting for the determinant of school opening decision becomes crucial to ensure a comparison between treated census areas, where the school year started earlier, and control areas, where the school year started later.

An important point for the research design depends on whether students go to schools close to their residence or not. Setting the distance threshold requires a clear understanding about the school-residence linkage. The traditional rule in Italy suggests sending the children to schools in the areas neighboring the residence, unless the household requires another location due to some particular situations, linked for instance to parents' job. This rule holds especially for kindergarten, primary and middle schools, where the subjects are equal across all

[15] See https://elezioni.comune.palermo.it/context.php?fc=1&tp=7. Similar situation in Catania with four secondary schools in a group of 395 stations. https://www.comune.catania.it/informazioni/servizi-eletorali/europee-2019/ubicazione-sezioni-elettorali/.

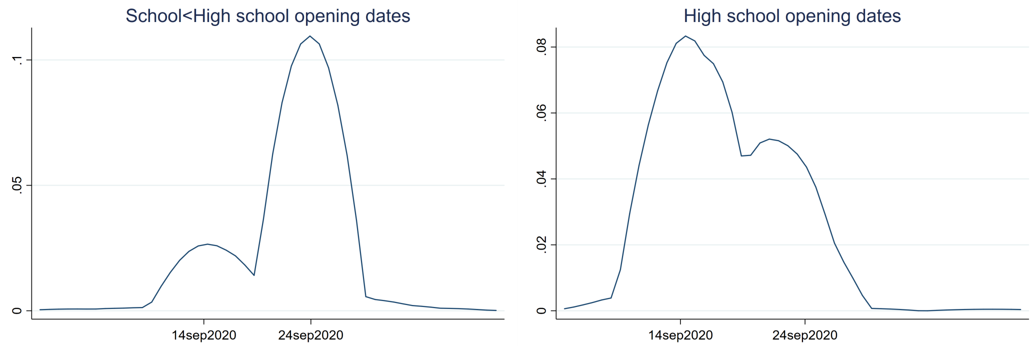


Figure 8: Date of first day in class (weighted by students) for level of schools

.

the schools. In contrast, secondary schools are different by the subject of education and thus determine much more mobility than the lower grades. According to official statistics, 79-83% of students younger than 15, thus including all the students from the infant up to middle school, employ less than 15 minutes to reach school from their residence. At the same time this percentage drops to 34 and 22% for the first and second cycle of high school (Istituto Nazionale di Statistica (ISTAT)).[16] Similar evidence emerges from the literature. A survey of Alietti et al. (2011) finds that 71-75% of students employ up to 10 minutes to go to school and reports that the vast majority of students attends primary school within 1 km from their house. This does not seem a peculiar Italian scenario. For Alberta's case in Canada, Bosetti and Pyryt (2007) highlight that 83% of parents send their children to their designated school, which is very close to their residence. Schneider et al. (1997) show that similar evidence holds for NYC districts, where 60% of the students in the district are accepted into their first-choice school. Overall, this suggests that the potential confounding effect deriving from students attending schools far away from their residence should not play a major role in Covid-19 diffusion, especially for early education levels.

To build our dummy indicator, a given census area is assigned to the average date of opening of the schools within 1 km of its ray weighted by the number of students. Figure 9 provides an example of the logic of this approach for a set of census areas and four schools in the city of Palermo. Census areas falling under the pink circle are assigned with the date of opening of the reference school. When a census area falls under more than one school, those ones in darker pink located between school three and school four, the resulting date will be the average date of those schools weighted by the number of students of each school. The empirical specification, then, integrates this information with a dummy activating two weeks after the weighted date of school opening in a given census area and remaining activated for the rest of the period, as typical of a Diff-in-Diff approach with time dummies and unit-level fixed effects, included also in this analysis. Most part of schools, indeed, had a progressive reopening with first year classes and 3 hours of lessons on the day of opening, and a gradual

[16] Another hint is given by the percentage of students that walk to go to the school that drops from 41% to 19 and 14% from the middle to the high schools Istituto Nazionale di Statistica (ISTAT).

reopening of the other classes during the next days. This means that direct effects should range between 1 and 16 days from reopening. The dummy segments the set of census areas in 26,925 units that get treatment across the time depending heterogeneously on the specific date on which school opened within their ray of 1 km, as in a Diff-in-Diff approach. Also, the data contains 9,756 cells that never get treated. Finally, he identification of the effect derives from the time discontinuity in treatment of the treated cells, accounting for the local level dynamic nature of Covid-19 pandemic evolution, both in other treated and controls cells. Being a Diff-in-Diff, the coefficient of the dummy has to be interpreted as a shift in the intercept of the model for the treated census areas.

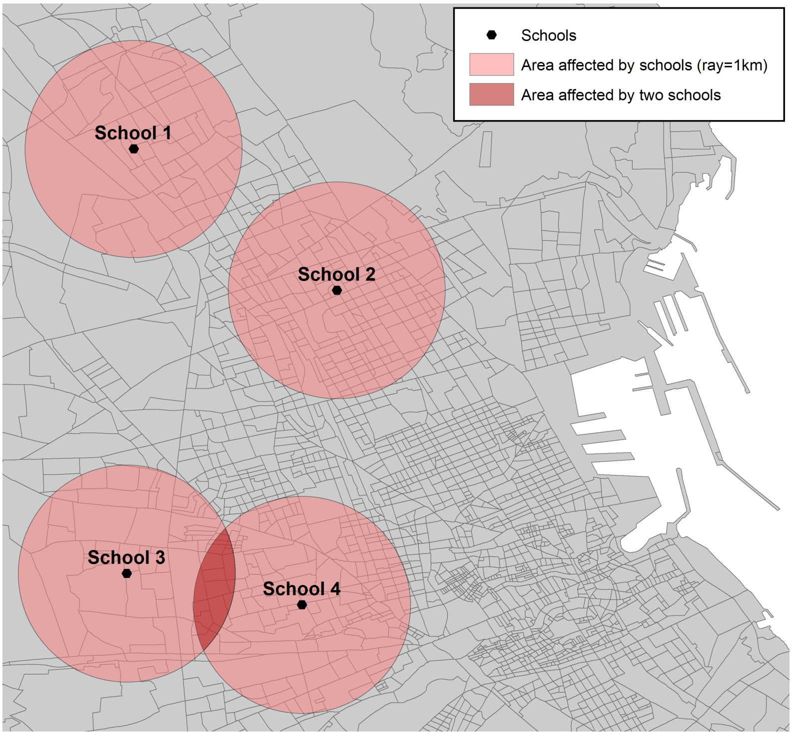


Figure 9: School opening effects on census areas

3.3 School variables and students coverage

To model the school-manager decision about the opening date dummy we described in the previous subsection, the analysis has complemented the information on date of opening with a set of indicators on the school level characteristics. These are obtained from the official school level data-sheet of the Ministry of Education, including average class size, number of teachers, number of non-permanent teachers, pupils with special needs

[17]. Some of these variables are present at a disaggregate school level, while others only at institute level (that is usually a group of more schools in the same area).

Table A1 in the Appendix reports summary statistics for individual school level variables as class size, number of students and pupils with special needs. These data show how secondary schools collect a relatively higher number of students, due to higher aggregation of education based on specific subjects. Table A1 reports also data for number of teachers and share of non-permanent teachers.

This analysis focuses only on public school because of two reasons. First, public schools were more likely to change their date following both the change in regulation from the Regional Government and the occurrence of the national referendum on September 20th-21st (see 2.3) because private schools are not used as polling stations. Second, the public school system includes and moves the vast majority of the school population in Sicily. A simple comparison is useful to see why understanding the potential role of schools in Covid-19 diffusion is relevant. As both Istituto Nazionale di Statistica (ISTAT) and Regional Department of Education data[18] suggest, the public schools hosted 717,524 students for the schooling year 2019/20, a number that increases up to 823,595 when considering teachers and other staff members, equal to about 16.5 percent of the total regional population or, just to give an idea, is equivalent to more than half of total employed in the region (Istituto Nazionale di Statistica (ISTAT))

In terms of the student population, ISTAT data for 2018/19 highlight that the percentage of students in public schools in Sicily is about 95.3%, with the lowest level in pre-primary (86.2%), where the school is not mandatory, and much higher in all other levels, equal to 96.5%, 99.0% and 95.3% in primary, middle and high school, respectively. On the other hand, the class size is larger: as we see in Table A1, the average class is slightly lower than 19 pupils (18.84), while the local unit average of employees in Sicily is 3.8 by the last census of 2011.[19]

[17]Data are extracted from the Ministry of Education data portal at the following link: https://dati.istruzione.it/espscu/index.html?area=anagScu

[18]Data from the Regional Department of education are extracted at the following link: https://www.usr.sicilia.it/index.php/dati-delle-scuole

[19]Pre-primary school include students from 3 to 6 years and is not mandatory, while others are mandatory and involve age classes of 6-11 for primary, 11-14 for middle and 14-19 for high schools.

Covid Economics 65, 20 January 2021: 47-77

Covid Economics 65, 20 January 2021: 47-77

Table 1: Share of students in public schools by provinces in 2018/19 (Source: ISTAT (2020a))

	Pre-primary	Primary	Middle	High	Total
Trapani	88.2%	99.4%	100.0%	98.4%	97.3%
Palermo	79.1%	94.5%	98.4%	94.7%	92.9%
Messina	87.3%	96.5%	98.5%	97.5%	95.7%
Agrigento	87.7%	99.0%	100.0%	96.4%	96.5%
Caltanissetta	89.4%	97.2%	100.0%	98.7%	97.0%
Enna	96.1%	100.0%	100.0%	98.0%	98.7%
Catania	89.0%	95.6%	98.4%	96.4%	95.3%
Ragusa	88.1%	96.7%	100.0%	96.4%	95.8%
Siracusa	87.7%	98.2%	99.5%	97.4%	96.4%
Sicily	86.2%	96.5%	99.0%	96.6%	95.3%

4 Empirical strategy

Consider the following dynamic panel regression model, which will be basis of our baseline specification

$$y_{i,t} = \alpha_i + \rho_1 y_{i,t-1} + \rho_2 y_{i,t-1} + \gamma t + \delta S_{i,t-2} + u_{i,t}, \quad i = 1, 2, \ldots, n, \qquad t = 1, 2, \cdots, T, \tag{1}$$

where $y_{i,t}$ denotes the natural logarithm of Covid-19 cases for census locality i at time t. The first two lagged values of this variable on the right-hand side are included to capture persistence in Covid-19 cases as implied by the Arellano-Bond test for autocorrelation to ensure that the errors serially uncorrelated over t. The key variable of interest is the dummy for school opening $S_{i,t-2}$ that takes the value 0 before the opening and 1 after the date on school opening. This variable enters the model with a lag of two weeks to reflect the serial time of infection and the delays in the detection of the virus among children. The parameter δ measures the causal effect of the opening of schools on Covid-19 cases. Additionally, the model includes the census areas fixed effects denoted by α_i that capture common shocks to the Covid-19 cases of all census localities t. $u_{i,t}$ is an error term capturing the remaining unobserved heterogeneity. Equation (1) is estimated using the Blundell-Bond two-step system GMM estimation method by allowing $S_{i,t-2}$ to be predetermined its lags as instruments.

While our estimation approach is able to account for time-invariant unobserved heterogeneity, time trend, one may still argue that the school opening decision is endogenous. Indeed, the decision of the school managers could be based on several reasons including the percentage of Covid-19 cases in the area, the preparation time due to longer management times for some internal organizational issue, and other administrative matters. This implies that an ideal estimation should account for this endogenous decision when modeling the effect of school opening. In doing so, the present analysis considers a two-stage problem, where the first-stage relates to the decision on the date of opening. In particular, the dummy for school opening $S_{it,}$ is modeled by a set of inverse probability weights, obtained from a Propensity Score Matching (PSM) estimation on several indicators affecting school opening decision for school managers.

In particular, the vector of indicators $\mathbf{z}_i$ includes the number of Covid-19 cases in August in the same census area, the average class size, the number of pupils with special needs, permanent and temporary teachers, the total number of schools within 1 km of ray. The algorithm employed for the matching is the five-nearest neighbor but the result is robust to other specifications such as kernel based matching or caliper matching. Formally, the PSM model is given by the probit model

$$e_i \stackrel{\text{def}}{=} \Pr(\bar{s}_i = 1|\mathbf{z}_i) = F(\mathbf{z}_i'\pi), \tag{2}$$

where $F(\cdot)$ is the Normal cumulative density function that models the probability of being treated on a set of determinants measured before the treatment. The treatment variable $\bar{s}_i$ for the PSM is a dummy activating whether the average date of school opening has been on the 14th of September or earlier, representing one of the two modes in Figure 7. We then obtain the propensity $\hat{\theta}_i$ that we use to build the weight $1/\hat{\theta}_i$ for the treated units and $1/(1-\hat{\theta}_i)$ for the control units. These weights are then incorporated in model (1) to weight both the dependent and the explanatory variables. Effectively, our approach involves a two-step estimation, where in the first step we obtain the propensity scores via model (2) and then in the second-step to estimate a weighted version of model (1). Table A4 in the Appendix introduces the results from the PSM, which are consistent with our expectations, e.g., the more the Covid-19 cases in August, the lower is the probability of opening earlier the schools. As the common support region in Figure 10 shows, for each treated census area, the PSM has found a counterfactual census area across all the score distribution, and therefore allows the inclusion of all the census areas in a weighted version of equation (1).

5 Results

This section presents our findings. Table 2 reports our main results of the baseline estimations on the whole sample of 33,366 inhabited census areas observed for 18 weeks. Columns 1 and 2 report the results based on the fixed effect estimator with one and two lags of the Covid-19 variable, respectively. Columns 3 to 5 report the results from the system GMM estimation, with Column 4 reporting the results based on the propensity score weights and Column 5 adding an interaction term between the school opening dummy and the second lag of the Covid-19 variable.

When not considering the interaction between school opening and the lagged level of Covid-19 in the census area, school opening increases the level of Covid-19 by between 1.5 and 2.9 percent. This result is robust to all the alternative specifications considered, including OLS-FE estimation not accounting for Nickell's bias in the

Covid Economics 65, 20 January 2021: 47-77

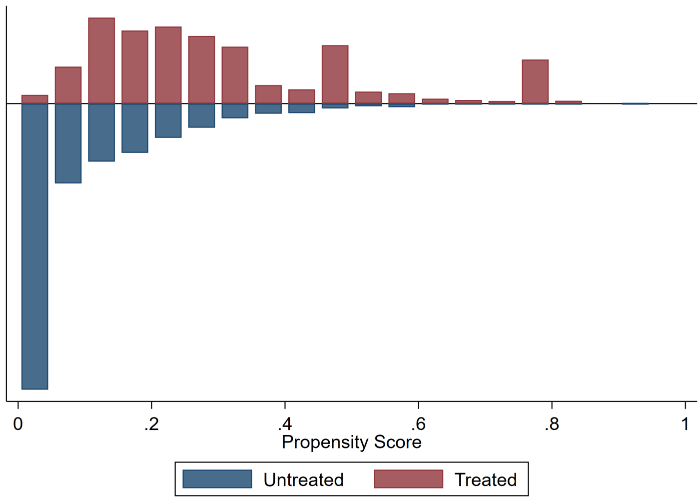


Figure 10: Common and out of support area of the Propensity Score Matching.

dynamic panel model in column 1-2, or when considering it as in column 3. The AR test suggests the presence of the first two lags of the lagged dependent variable, while the Hansen test on over-identifying restriction excludes over-identification at 5 percent of probability level. Column 4 reports the estimates obtained when weighting equation (1) with inverse propensity score weights built on the determinant of school opening decisions, which remain consistent to what was found in the previous three models, with an average increase in cases of about 1.9 percent associated to school opening. Also, accounting for the endogenous decision about school opening improves the Hansen test's performance, suggesting that the orthogonality conditions are valid at the required level of probability. Finally, results in column 5 include the interaction between the dummy on school opening and the level of Covid-19 cases at the time of the opening of schools, to account for the dynamic role of school opening on the diffusion. Again, also, in this model the estimates remain positive and consistent. Besides, the coefficient associated with the interaction term is positive and significant, suggesting a dynamic additional increase in cases due to school opening. Note how this effect is relatively lower than the influenza reduction effect of Ali et al. (2018) given by school closures in Hong Kong. This is an interesting comparison due to the same time window, that is, the weekly average coefficient. The size of difference (60% concerning our favorite benchmark of table 2) may be explained by considerable heterogeneity of type of virus transmission, period, data, and use of masks.

5.1 Heterogeneity analysis

How may the above effect change depend on the school and population characteristics? This subsection investigates this question using a set of cross-sectional variables available at census-area level. Table 3 reports

Covid Economics 65, 20 January 2021: 47-77

Table 2: Covid-19 cases and school opening: GMM estimations

Dependent variable: Covid-19 cases (ln)	(1)	(2)	(3)	(4)	(5)
School opening (1=yes, 2nd lag)	0.029***	0.026***	0.015***	0.019***	0.017***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
School opening (1=yes, 2nd lag) X Covid-19 (ln, 2nd lag)					0.080**
					(0.033)
Covid-19 (ln, lag)	0.173***	0.155***	0.795***	0.808***	0.799***
	(0.004)	(0.004)	(0.020)	(0.032)	(0.031)
Covid-19 (ln, 2nd lag)		0.094***	0.029***	0.029***	-0.038
		(0.003)	(0.007)	(0.011)	(0.030)
Time dummies	Y	Y	Y	Y	Y
Unobserved time invariant heterogeneity	Y	Y	Y	Y	Y
Observations	597,619	597,619	597,619	597,619	597,619
R-squared	0.093	0.101			
Number of Census areas	33,366	33,366	33,366	33,366	33,366
Arellano-Bond test for AR(1) in first differences			0.000	0.000	0.000
Arellano-Bond test for AR(2) in first differences			0.000	0.000	0.000
Arellano-Bond test for AR(3) in first differences			0.150	0.243	0.292
Hansen p-value			0.051	0.122	0.404
Model	OLS-FE	OLS-FE	SYS-GMM	SYS-GMM	SYS-GMM
Propensity scores weighting	N	N	N	Y	Y

Two-step robust standard errors are in parentheses. Orthogonal difference transformation applied. Census cells fixed effects were included in all regressions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

three types of results. Column 1 estimates the effect of reopening on cases linked only to people directly involved in schools as students, teachers, staff. While the effect is strongly significant, the point estimate is much lower, and the dynamic process loses its time persistency. At least two arguments may explain the lower magnitude and the loss of persistency observed. First, as already introduced above, the school population mostly involves youths, which may be more likely to be asymptomatic. Thus, the observed increase in cases is lower than the real increase. Second, schools may have been efficient in isolating classes with positive cases, a condition that may explain both the lower increase in cases and the absence of a dynamic within the school population. Matching this result with what was observed in table 2, this suggests that most of the direct contagion within the school system, then develops in other contexts, such as within the families other social networks. In this sense, schools may act as the initial spark of a larger contagion within these networks.

A second essential type of heterogeneity derives from the kind of school. To study this mechanism, the analysis is conducted separately for schools lower than the high schools, including infancy, primary and middle school, and for the high school itself. As explained above, this clustering is justified by the different spatial mobility students show concerning the type of school. Columns 2 and 3 report the result obtained when activating the school opening dummy separately for these two groups. As expected, the coefficient associate with infancy, primary and middle schools is much higher and more significant than the one associated with middle schools. On average, the first one is related to an increase in the census-area case by 1.2 percent, while

Covid Economics 65, 20 January 2021: 47-77

the second ones are linked to a much lower increase, equal to about 0.04 percent. It is important to underscore again how the difference in this result is strongly linked to the difference in the students' local spatial dynamic, which are more likely to go to schools within a ray of 1 km when attending school lower than middle school. It is, therefore, possible that the high schools may have a similar effect but more spatially dispersed across the territory of the municipality, a result that is still compatible with Munday et al. (2020), who suggest that secondary schools may be more robust drivers of contagion.

Finally, columns 4 and 5 introduce the heterogeneity results across class size, conducted clustering the regression for census areas with average class size below and above the median value, equal to about 20 students. In this case, the estimated effect appears the effect of school opening is not different from zero for schools with smaller size classes. At the same time, it is equal to +2.1% in areas with average classes larger than the median. This is in line with recommendations of Lordan et al. (2020) and suggest that reducing the number of students per class may reduce the contagion induced by school opening.

Covid Economics 65, 20 January 2021: 47-77

Table 3: Covid-19 cases and school opening. Heterogeneity linked to schools

	(1)	(2)	(3)	(4)	(5)
	Covid-19 within school	Below high school	High school	Class size≤median	Class size>median
School opening (2nd lag)	0.003*** (0.001)			0.070 (0.132)	0.021*** (0.003)
School<sec. opening (2nd lag)		0.012*** (0.004)			
School sec. opening (2nd lag)			0.004* (0.002)		
School opening (1=yes) X 2nd lag of the dep. var.	-0.283 (0.231)	0.088*** (0.033)	0.107*** (0.032)	1.340** (0.600)	0.065* (0.034)
Covid-19 within schools (ln, lag)	0.063 (0.174)				
Covid-19 within schools (ln, 2nd lag)	0.255 (0.194)				
Covid-19 (ln, lag)		0.801*** (0.031)	0.801*** (0.031)	0.708*** (0.060)	0.816*** (0.038)
Covid-19 (ln, 2nd lag)		-0.045 (0.030)	-0.061** (0.029)	-1.246** (0.569)	-0.021 (0.031)
Time dummies	Y	Y	Y	Y	Y
Unobserved time invariant heterogeneity	Y	Y	Y	Y	Y
Obs	597.619	597.619	597,619	208,887	388,732
Number of Census areas	33,366	33,366	33,366	11,659	21,707
Arellano-Bond test for AR(1) in first differences	0.003	0.000	0.000	0.000	0.000
Arellano-Bond test for AR(2) in first differences	0.918	0.000	0.000	0.000	0.000
Arellano-Bond test for AR(3) in first differences	0.931	0.285	0.298	0.901	0.298
Hansen p-value	0.164	0.319	0.402	0.658	0.610

Two-step robust standard errors are in parentheses. All dependent variables are in natural logarithm. All regressions are weighted by propensity scores. Orthogonal difference transformation applied. Census cells FE were included in all regressions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Table 4 reports the heterogeneous analysis across the population characteristics. Columns 1 and 2 consider the dependent variable for only the population above the school-age (older than 19) or within the school age. As expected, the impact appears higher in magnitude for individuals outside the school age. An estimated increase equal to 1.5 percent, more than double of the one estimated for individuals within school age, equals about 0.6

percent. As for the results about the school population in Table 3, even in the younger population, it appears less persistent. Finally, the last two columns' results show that the effect of school opening is significantly stronger in less populated areas. While this result may be unexpected, a potential explanation certainly relies on the fact that, in these areas, schools act as a social collector and potentially represent a higher share of interaction concerning schools in big cities.

Table 4: Covid-19 cases and school opening. Heterogeneity linked to population

	Covid-19 cases≤19 yo. (1)	Covid-19 cases>19 yo. (2)	Covid-19 cases pop density≤median (3)	Covid-19 cases pop density>median (4)
School opening (2nd lag)	0.006***	0.015***	0.014***	0.005
	(0.001)	(0.003)	(0.002)	(0.008)
School opening (2nd lag) X 2nd lag dependent	-0.040	0.054	0.031	0.122**
	(0.117)	(0.037)	(0.029)	(0.059)
Covid-19 ≤ 19 yo (ln, lag)	0.758***			
	(0.105)			
Covid-19 ≤ 19 yo (ln, 2nd lag)	0.049			
	(0.106)			
Covid-19 >19 yo (ln, lag)	0.797***			
	(0.032)			
Covid-19 >19 yo (ln, 2nd lag)		-0.020		
		(0.033)		
Covid-19 (ln, lag)			0.814***	0.771***
			(0.029)	(0.028)
Covid-19 (ln, 2nd lag)			0.002	-0.074
			(0.018)	(0.054)
Time dummies	Y	Y	Y	Y
Unobserved time invariant heterogeneity	Y	Y	Y	Y
Observations	597,619	597,619	298,766	298,853
Number of Census areas	33,366	33,366	16,668	16,698
AR1 probability	0.000	0.000	0.000	0.000
AR2 probability	0.000	0.000	0.000	0.000
AR3 probability	0.475	0.346	0.769	0.120
Hansen Probability	0.180	0.374	0.253	0.114

Two-step robust standard errors are in parentheses. All dependent variables are in natural logarithm. All regressions are weighted by propensity scores. Orthogonal difference transformation applied. Census cells fixed effects were included in all regressions. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

5.2 Dynamic accounting analysis

As a consequence of the above estimates, one may ask what would have been the Covid-19 cases if the school would not have opened. Our model's dynamic nature allows us to predict a set of counterfactual scenarios where social interactions linked to school would have been lower or zero, with zero denoting full school closure. This exercise's results need to be taken with caution, as the present work may make only a few hypotheses about how social interaction may have changed without school opening. However, it is impossible to predict whether the interaction between pupils outside school would have increased due to school closure. For example, parents could have felt safer to allow their children to go out, reducing the effect of school closure.

For the sake of this exercise, we use the most conservative estimate obtained predicting the model employed for the results in column 2 of Table 4. This allows us to exclude the effect of high schools, which are more spatially dispersed and closed in mid-November. The analysis creates three scenarios, the first scenario where the prediction is calculated, setting the dummy and the interaction term at 0, involving zero pupil interactions and complete school closure. The second and third scenarios reduce by 75 and 50 percent the social interactions,

Covid Economics 65, 20 January 2021: 47-77

setting, therefore, the dummy and the interaction term to 0.25 and 0.50. This situation may result from an alternation of distance learning with face-to-face learning or through the implementation of other precautionary measures.

Figure 11 suggests the number of predicted Covid-19 cases would have shown different dynamics and lower numbers with the three different scenarios above. In particular, the total number of patients would have decreased to a value between 52,406 and 57,676 from the observed 66,092 Covid-19 cases for the period under consideration. In relative terms, these estimates translate into a cumulative increase of between 14.6 and 26.1 percent of cumulative Covid-19 cases calculated on December 14th. The magnitude of the decrease depends on whether the school closures would have entailed the absence of any other social contact of the students with their school social network, or whether the reduction of contacts would have been just equal to 50 percent of the baseline.

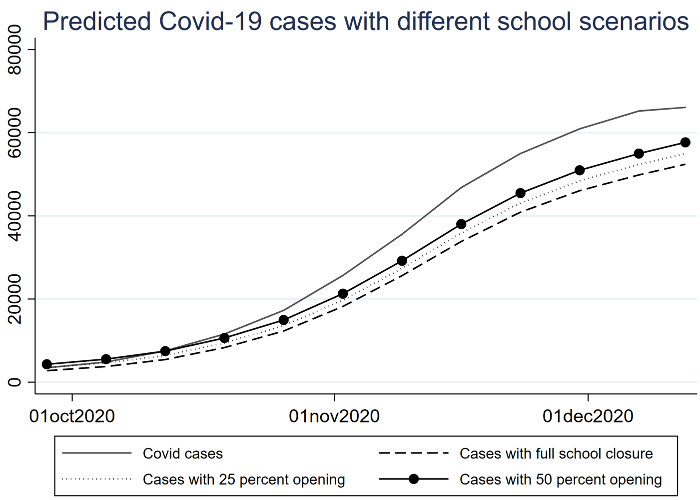


Figure 11: Cumulative Covid-19 cases and predicted cases with three school closure scenarios.

5.3 Robustness

This section introduces the results of a set of exercises to test the robustness of the main results. The first test consists of switching the time of observation from the serial time, equal to 7 days, to the incubation period of 5 days (Cereda et al. (2020)). This exercise requires also changing the time lag of the dummy school lag from 2 to 3 to consider the same period of days.[20]. Column 1 of Table 5 reports the estimated coefficient, which is slightly smaller (1%) in magnitude but still strongly significant, with the Covid-19 autoregressive coefficients remaining unchanged.

A second robustness test consists of replacing the dependent variable's natural log with the correspondent

[20]In this way, we believe still 15 days, given by three lags times five days instead of 14, given by two lags times one week

Covid Economics 65, 20 January 2021: 47-77

inverse hyperbolic transformed value, following the methodology developed by Bellemare and Wichman (2020). As column 2 reports, the estimated coefficients remain positive and significant, with a level of a magnitude consistent with what is displayed in Table 2.

As the third robustness test, we run the baseline model substituting the dependent variable with the share of Covid-19 cases over the total census population. As already discussed, given that precise population estimates at the census area level are available only for 2011, calculating the share is prone to error due to migration and residence changes. According to official estimates, about 738,000 individuals have changed residence between 2011 and 2018 (Istituto Nazionale di Statistica (ISTAT))[21] in Sicily. When projecting this share to the ten years, this could have involved up to 21% of the population of the island population, which still represents a lower bound because some people moved their residence without formally communicating it to the authorities. These figures justify why using population shares would have been quite unreliable. However, testing this with the available data could remain essential to understanding whether the effect is somehow driven by population level in a given census area. As column 3 shows, the result remains consistent when adopting the share of population positive to Covid-19 as a dependent and autoregressive variable. The estimated increase in level is equal to about 1.5% and strongly significant, a result that remains in line with the benchmark.[22].

As a final test, the analysis is conducted a reduced time period, including observations for only one month before the first mode of school opening (September 14th) and one month after the second mode (September 24th). Results from the reduced period, displayed in column 4, are still strongly positive and significant. Even the auto-regressive tests suggest adding lag[23].

Covid Economics 65, 20 January 2021: 47-77

[21]Unfortunately ISTAT does not report the changes of residence for 2019 and 2020.
[22]The analysis used first differenced data for Covid-19 as level equation report a unit root of 1.04, results available under request
[23]Specification with three lags solves Hansen and AR test problems, and the coefficient of schooling is still 1.5%. Results available under request.

Covid Economics 65, 20 January 2021: 47-77

Table 5: Covid-19 cases and school opening. Robustness

	5 days periods	Hyperbolic transformed value	Cases over Pop 2011	Reduced time span
Dependent variables:	Covid-19 cases (ln)	Covid-19 cases (IHS)	Cases over population	Covid-19 cases (ln)
	(1)	(2)	(3)	(4)
School opening (2nd lag)	0.010***	0.022***	0.015***	0.015***
	(0.002)	(0.004)	(0.002)	(0.003)
School opening (2nd lag) X 2nd lag dependent	0.016	0.081**	0.031	0.118***
	(0.013)	(0.033)	(0.029)	(0.037)
Dependent variable (lag)	0.780***		-0.001	0.864***
	(0.042)		(0.015)	(0.033)
Dependent variable (2nd lag)	0.050***		0.033**	-0.081**
	(0.011)		(0.015)	(0.033)
Covid-19 IHS (lag)		0.799***		
		(0.031)		
Covid-19 IHS (2nd lag)		-0.039		
		(0.030)		
Time dummies	Y	Y	Y	Y
Unobserved time invariant heterogeneity	Y	Y	Y	Y
Observations	829,922	597,619	537,472	531,227
Number of Census areas	33,366	33,366	31,616	33,356
AR1 probability	0.000	0.000	0.000	0.000
AR2 probability	0.000	0.000	0.000	0.000
AR3 probability	0.913	0.292	0.327	0.038
Hansen Probability	0.245	0.395	0.569	0.01

Two-step robust standard errors are in parentheses. All dependent variables are in natural logarithm. All regressions are weighted by propensity scores. Orthogonal difference transformation applied. Census cells fixed effects were included in all regressions. Statistical significance: ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

CEPR PRESS

6 Conclusions

This work has employed a design based on differences in school opening time to test for the localized total effect of opening a school concerning Covid-19 increase in the census area. The dataset includes more than 66,000 geo-localized Covid-19 cases for the Sicily region, matched with precise information on the date on which schools have opened in the surrounding area. The time discontinuity of school opening derives from a change in the regulation that occurred two weeks before the school year's official start, together with a referendum that has delayed the opening of schools selected as seats of poll stations. These two conditions have generated a wide heterogeneity on the date of commencement. The endogeneity of school manager decisions has been modeled as a two stage problem. Therefore, the empirical strategy has involved a propensity score model in weighting the census areas on the school-manager determinants of school opening. Thus, the selected estimating strategy consisted of a weighted Blundell-Bond estimation, able to account for the dynamic evolution of Covid-19 infection in a given census area. To the best of our knowledge, this is the first work accounting for the dynamic process of Covid-19. Unlike the previous literature, which has based its estimates on regional or province data, this work is the first to rely on very granular geocoded data, measured at census area, which corresponds to about a block of 0.13 km^2. This work can also investigate the impact of school opening both on cases that occurred within the schools and on claims that occurred in the geographical areas where the students reside. In this sense, the estimates obtained from this exercise can be considered the global (direct and indirect) localized effects of school opening.

Results show that nearby schools observed a positive short-run localized increase of +1.5-2.9% in the Covid-19 cases after the school opening. Besides, school opening appears to play an additional role by increasing non-linearly the number of instances in zones hosting some Covid-19 cases. Using the estimated coefficients, the analysis makes a dynamic accounting of different school opening scenarios, suggesting the final increase in Covid-19 cases associated with school may have been equal to between 14.6 and 26.1 percent of the cumulative Covid-19 cases observed on December 14th.

Finally, a set of potential mechanisms and policy options emerge from the heterogeneity tests presented in section 4.3. First, larger class sizes are associated with a higher impact of school opening on Covid-19, while reducing the number of students per class appears to reduce infection potential. Second, even though school opening involves most of the region's youth population, the impact on Covid-19 cases is more substantial for a population older than 19. This may reflect that many students remain asymptomatic and may spread their infection in families or social networks outside the school. Increasing the number of testing within the schools remain, therefore, crucial to reduce the disease. Finally, the contagion appears to be higher in zones more sparsely populated, highlighting the relevance of stronger social interactions.

Covid Economics 65, 20 January 2021: 47-77

References

Agostinelli, F., M. Doepke, G. Sorrenti, and F. Zilibotti, 2020, When the Great Equalizer Shuts Down: Schools, Peers, and Parents in Pandemic Times, NBER Working Paper No. 28264.

Ali, S.T., B.J. Cowling, E.H.Y. Lau, V.J. Fang, and G.M. Leung, 2018, Mitigation of Influenza B Epidemic with School Closures, Hong Kong, 2018, *Emerging Infectious Diseases* 24, 2071–2073.

Alietti, A., D. Renzi, M. Vercesi, and A. Prisco, 2011, Children's Independent Mobility in Italy, .

Bellemare, M. F., and C. J. Wichman, 2020, Elasticities and the inverse hyperbolic sine transformation, *Oxford Bulletin of Economics and Statistics* 82, 50–61.

Bosetti, L., and M. C. Pyryt, 2007, Parental Motivation in School Choice: Seeking the Competitive Edge, *Journal of School Choice* 1, 89–108.

Buonsenso, D., C. De Rose, R. Moroni, and P. Valentini, 2020, SARS-CoV-2 infections in Italian schools: preliminary findings after one month of school opening during the second wave of the pandemic, *medRxiv - Pediatrics*.

Carlson, C. J., A. C. R. Gomez, S. Bansal, and S. J. Ryan, 2020, Misconceptions about weather and seasonality must not misguide COVID-19 response, *Nature Communications* 11.

Cauchemez, S., A.J. Valleron, P.Y. Boelle, A. Flahault, and N.M. Ferguson, 2009, Estimating the impact of school closure on influenza transmission from Sentinel data, *Lancet Infection Disease* 9, 473–481.

Cereda, D., M. Tirani, F. Rovida, V. Demicheli, M. Ajelli, P. Poletti, F. Trentini, G. Guzzetta, V. Marziano, A. Barone, M. Magoni, S. Deandrea, G. Diurno, M. Lombardo, M. Faccini, A. Pan, R. Bruno, E. Pariani, G. Grasselli, A. Piatti, M. Gramegna, F. Baldanti, A. Melegaro, and S. Merler, 2020, The early phase of the COVID-19 outbreak in Lombardy, Italy, .

Chudik, A., K. Mohaddes, M. H. Pesaran, M. Raissi, and A. Rebucci, 2020, A Counterfactual Economic Analysis of Covid-19 Using a Threshold Augmented Multi-Country Model, NBER Working Paper No. 27855.

Engzell, P., A. Freyd, and M. Verhagen, 2020, Learning inequality during the COVID-19 pandemic, .

Gandini, S., M. Rainisio, M.L. Iannuzzo, F. Bellerba, F. Cecconi, and L. Scorrano, 2020, No evidence of association between schools and SARS-CoV-2 second wave in Italy, .

Haug, N., L. Geyrhofer, A. Londei, E. Dervic, A. Desvars-Larrive, V. Loreto, B. Pinior, S. Thurner, and P. Klimek, 2020, Ranking the effectiveness of worldwide COVID-19 government interventions, *Nature Human Behaviour* 4, 1303–1312.

Covid Economics 65, 20 January 2021: 47-77

Hsiang, S., D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L.Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, and T. Wu, 2020, The effect of large-scale anti-contagion policies on the COVID-19 pandemic, *Nature* 584, 262–285.

Isphording, I.E., M. Lipfert, and N. Pestel, 2020, School Re-Openings after Summer Breaks in Germany Did Not Increase SARS-CoV-2 Cases, IZA DP No. 13790.

Istituto Nazionale di Statistica (ISTAT), 2020, Dati su Istruzione e Formazione Scolastica.

———, 2020b, Dati su Popolazione residente al 1 Gennaio.

———, 2020c, Indagine Multiscopo sulle famiglie: aspetti della vita quotidiana.

———, 2020d, Popolazione e famiglie. Migrazioni: Trasferimenti di residenza.

Istituto Superiore di Sanità (ISS), 2020, Dati indice riproduzione.

Kawano, S., and M. Kakehashi, 2015, Substantial Impact of School Closure on the Transmission Dynamics during the Pandemic Flu H1N1-2009 in Oita, Japan, *PLoS One* 10, 1–15.

Lattanzio, S., 2020, La scuola un focolaio?, lavoce.info.

Lewis, D., 2020, Why Schools probably aren't Covid Hotspots, *Nature* 587, 17–17.

Li, X., W. Xu, M. Dozier, Y. He, A. Kirolos, and E. Theodoratou, 2020, The role of children in transmission of sars-cov-2: A rapid review, *Journal of Global Health* 10, 1.

Lordan, R., G.A. Fitzgerald, and T. Grosser, 2020, Reopening schools during COVID-19, *Science* 369, 1146–1146.

Maltezou, H.C., I. Magaziotou, X. Dedoukou, E. Eleftheriou, V. Raftopoulos, A. Michos, A. Lourida, M. Panopoulou, K. Z. Stamoulis, V. Papaevangelou, E. Petinaki, A. Mentis, A. Papa, A. Tsakris, E. Roilides, G.A. Syrogiannopoulos, and M. Tsolia, 2020, Children and adolescents with sars-cov-2 infection, *The Pediatric Infectious Disease Journal* 39, e388–e392.

McAloon, C., A. Collins, K. Hunt, A. Barber, A.W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, P. Wall, M. Green, L. O'Grady, and S.J. More, 2020, Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research, *BMJ Open* 10, e039652.

Mehta, N.S., O.T. Mytton, E. W. S. Mullins, T.A. Fowler, O.B. Falconer, C.L. andMurphy, C. Langerberg, W.J.P. Jayatunga, D.H. Eddy, and J.S. Nguyen Van Tam, 2020, Sars-cov-2 (covid-19): What do we know about children? a systematic review, *Clinical Infectious Diseases* 71, 2469.

Covid Economics 65, 20 January 2021: 47-77

Munday, J.D., K. Sherratt, S. Meakin, A. Endo, C. A. B. Pearson, J. Hellewell, S. Abbott, N. Bosse, K. A. Atkins, J. Wallinga, W. J. Edmunds, A. Jan van Hoek, and S. Funk, 2020, Implications of the school-household network structure on SARS-CoV-2 transmission under different school reopening strategies in England, *medRxiv*.

Oster, E., 2020, National COVID-19 School Response Dashboard, .

Psacharopoulos, G., V. Collis, H.A. Patrinos, and E. Vegas, 2020, Lost Wages: The COVID-19 Cost of School Closures, IZA DP No. 13641.

Schneider, M., P. Teske, M. Marschall, M. Mintrom, and C. Roch, 1997, Institutional Arrangements and the Creation of Social Capital: The Effects of Public School Choice, *American Political Science Review* 91, 82–93.

Sebastiani, G., and G. Palú, 2020, COVID-19 and School Activities in Italy, *Viruses* 12.

Tatem, A. J., 2017, WorldPop, open data for spatial demography, *Scientific data* 4, 1–4.

Covid Economics 65, 20 January 2021: 47-77

Other data specifications and results

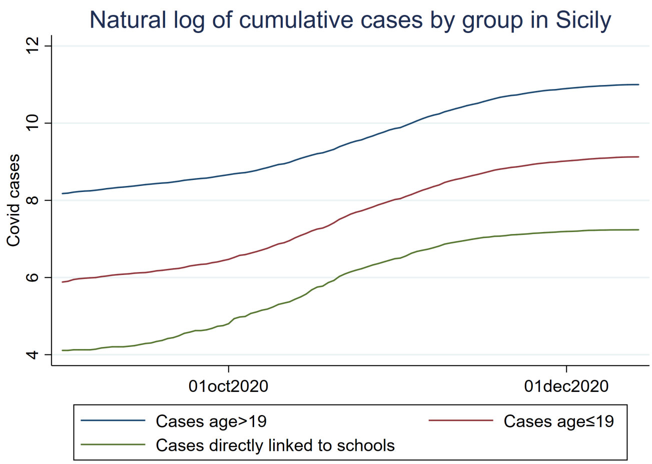


Figure A1: Natural log of cumulative cases by group from September onwards.

Covid Economics 65, 20 January 2021: 47-77

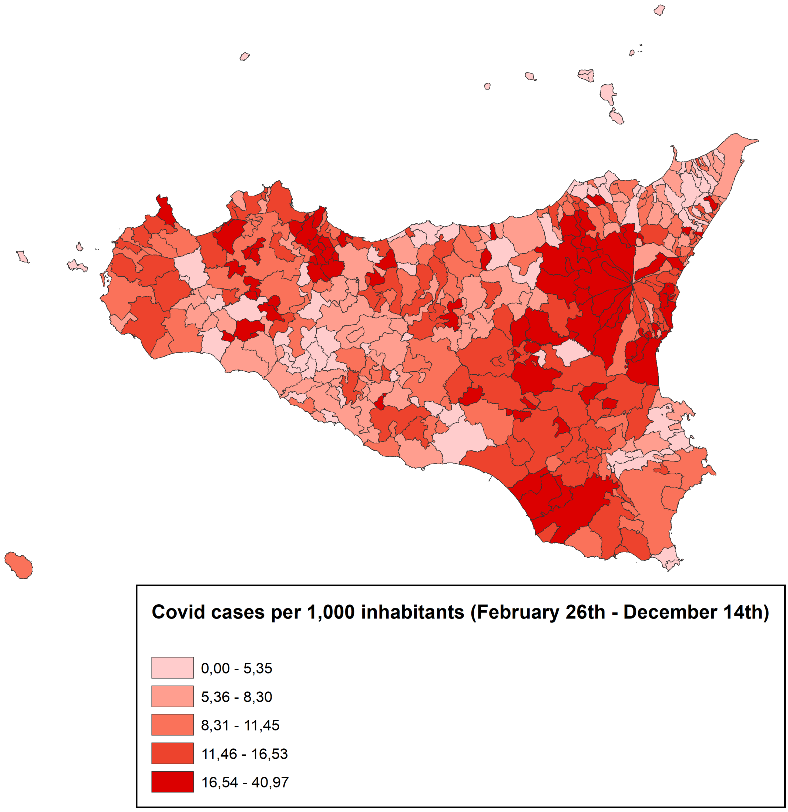


Figure A2: Covid-19 cases by population at 2011 whole period

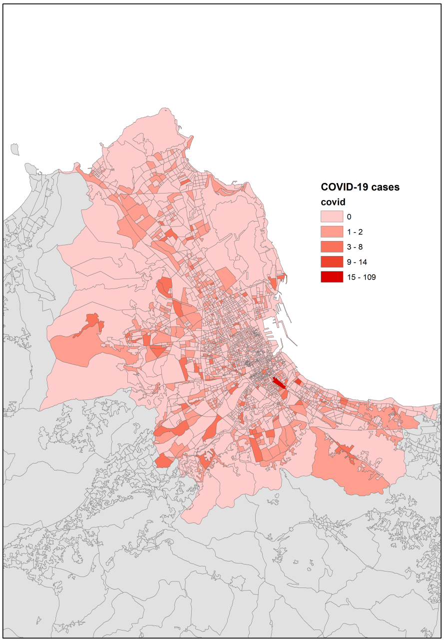


Figure A3: Cumulative cases by population at 2011 at October 15th: census areas of Palermo

Table A1: Public school characteristics in Sicily

School type	N. of pupils	class size	% of pupils with special needs	Teachers	% of temporary teachers
pre-primary					
Mean	69.1	19.93	2.6%		
Sd. dev.	47.26	3.97	0.027		
Min	3	3	0		
Max	342	39	26.7%		
primary					
Mean	158.1	17.05	4.8%		
Sd. dev.	119.5	3.74	0.037		
Min	3	3	0		
Max	798	37.5	33.8%		
middle					
Mean	232.5	18.70	5.2%		
Sd. dev.	172.4	3.47	0.038		
Min	8	8	0		
Max	978	31.3	29.4%		
secondary					
Mean	360.9	20.20	3.6%		
Sd. dev.	379.4	4.34	0.044		
Min	5	5	0		
Max	2688	40	27.5%		
all levels					
Mean	169.9	18.84	3.7%	103.1	0.12
Sd. dev.	207.8	4.11	0.037	33.76	0.085
Min	3	3	0	2	0
Max	2688	40	33.8%	344	0.9

Covid Economics 65, 20 January 2021: 47-77

CEPR PRESS

Table A2: Summary Statistics of the main variables used for the analysis

Variables	N	mean	sd	min	max
Covid-19 cases	733,620	0.09	0.52	0	75
Covid-19 cases above 19 yo	733,620	0.08	0.46	0	73
Covid-19 cases younger or 19 yo	733,620	0.01	0.14	0	21
Covid-19 within school	733,620	0.00	0.05	0	4
School opening (1=yes)	733,620	0.44	0.50	0	1
Population density	36,661	137.91	231.84	0	3036

Table A3: Rt Weekly reproduction rate for Sicily

Reference week	Rt	lower bound	upper bound
August 24-30	0.82	0.53	1.16
August 31 - September 6	0.96	0.66	1.32
September 21-27	1.08	0.70	1.43
September 14-20	1.2	0.77	1.75
September 21-27	1.19	0.88	1.57
September 28 - October 4	1.22	0.79	1.64
October 5-11	1.23	0.88	1.69
October 12-18	1.28	0.99	1.49
October 19-25	1.42	1.21	1.61
October 26 - November 1	1.4	1.15	1.69
November 2-8	1.18	1.02	1.44
November 9-15	1.13	0.95	1.26
November 16-22	1.05	0.73	1.31
November 23-29	0.84	0.62	1.16
November 30 - December 6	0.72	0.59	0.88
December 7 - 13	0.73	0.64	0.82
December 14-20	0.8	0.71	1.01

Table A4: Scores from the Propensity Score Matching

	Coef.	Std. Err.
Covid-19 cases in August	-0.003*	(0.002)
Class size	0.102***	(0.004)
Pupils with special needs	14.48***	(0.689)
Teachers	0.001***	(0.000)
Temporary Teachers	0.321***	(0.031)
Number of Schools	-0.053***	(0.001)
Constant	-3.220***	(0.083)
Obs		36681
Pseudo R^2		0.27

Robust standard errors are in parentheses. Dependent variable is treatment date at 14 September. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

Covid Economics 65, 20 January 2021: 47-77

Could the United States benefit from a lockdown? A cost-benefit analysis[1]

Anna Scherbina[2]

Date submitted: 12 January 2021; Date accepted: 14 January 2021

Though COVID vaccines are finally available, the rate at which they are administered is slow, and in the meantime the pandemic continues to claim about as many lives every day as the 9/11 tragedy. I estimate that with the promised rate of vaccinations, if no additional non-pharmaceutical interventions are implemented, 406 thousand additional lives will be lost and the future cost of the pandemic will reach $2.4 trillion, or 11% of GDP. Using a cost-benefit analysis, I assess whether it is optimal for the United States to follow the lead of many European countries and introduce a nation-wide lockdown. I find that a lockdown would be indeed optimal and, depending on the assumptions, it should last between two and four weeks and will generate a net benefit of up to $1.2 trillion.

Covid Economics 65, 20 January 2021: 78-107

1 I am grateful to Dan Bergstresser, Steve Cecchetti, Kevin Corinth, Josh Goodman, James Ji, Bob Kaplan, Joel Lander, David Levine, Jean-Paul L'Huillier, Peter Limbach, Bob McDonald, Debarshi Nandy, Andreas Neuhierl, Peter Petri, Bernd Schlusche and seminar participants at George Mason University, Brandeis University and The American Enterprise Institute for very helpful comments and suggestions.

2 Associate Professor, Brandeis International Business School, Brandeis University.

Covid Economics 65, 20 January 2021: 78-107

I. Introduction

Operation Warp Speed has successfully delivered two highly effective COVID-19 vaccines, with additional vaccine candidates undergoing clinical trials. However, vaccine production and distribution are slow, with the vaccination target of 70% expected to be reached only by the end of May 2021 according to more optimistic estimates.[1] In the meantime, the number of new infections is at peak levels, and the virus claims about as many lives every day as the 9/11 tragedy.

While some European countries began a new round of national lockdowns, there is resistance to implementing more stringent COVID restrictions in the United States.[2] The costs of a lockdown are felt in real time in the form of inconveniences and lost wages while the benefits from the reduced number of illnesses and deaths come in the future, and as such they may be perceived as hypothetical and underestimated. Moreover, the public may view the pandemic risks as acceptable because children are largely unaffected and because frontline workers and first responders getting protection from the virus by being among the first to be vaccinated (e.g., Tumpey et al. (2018), Table 12.1).

Despite society as a whole being potentially less concerned about saving the lives of the more vulnerable older adults,[3] the older people's lives are valuable to them.[4] The value of life can be quantified by a person's willingness to pay to stay alive, with metrics such as the value of statistical life (VSL) and discounted quality-adjusted life years (dQALY) being widely used in policy decisions. Moreover, the fatality data shows that COVID-19 also poses substantial risks to the lives of younger people who may be unaware of their health vulnerabilities ex-ante and therefore fail to take adequate precautions.

The COVID experience from around the world has shown that centralized policies are critical to achieving an optimal pandemic management. The failed Swedish experiment has illustrated

[1] See, e.g, https://www.technologyreview.com/2020/12/01/1012817/us-official-says-every-american-who-wants-a-covid-19-vaccine-will-have-one-by-june/.

[2] I will use the terms "COVID" and "COVID-19" interchangeably.

[3] See, e.g., https://www.texastribune.org/2020/04/21/texas-dan-patrick-economy-coronavirus/.

[4] For example, a Gallup poll showed that older people were more willing than younger people to choose resuscitation or ventilator support when asked about preferences in the event of terminal illness (Gallup and Newport (1991)).

that it may be impossible to selectively protect the vulnerable population without a government intervention.[5] Analysing U.S. data, Boehmer et al. (2020) find that increased rates of infection among young people in the June–August 2020 period helped transmit the virus to more vulnerable high-risk groups, such as older adults. This happened in spite of the broad awareness of higher risks faced by the older population.

Even when a COVID infection is not fatal, it is still costly because the sick consume medical services that could have been allocated to other health conditions. They also miss days of productive work, reducing the GDP (or in the case of children and older adults, their caretakers miss productive work days). I perform a cost-benefit analysis of a possible lockdown by comparing its benefits that come from reducing the number of future infections until the vaccination target is reached to the incremental costs it would impose on the economy and finding the optimal stopping time before incremental costs start to exceed incremental benefits. I model the COVID-19 pandemic curve using the SIR (susceptible, infected, recovered) model widely used in epidemiology. I use estimates from the COVID literature to obtain the model parameters, such as the basic reproduction number that prevails with the social distancing measures currently in place, as well as estimates of what it will be with a nation-wide lockdown, similar to lockdowns implemented in Europe in Spring 2020.

The expected future monetary cost of the COVID pandemic is calculated from the following three components: (1) the loss of productivity due to missed work of the symptomatically ill; (2) the cost of medical interventions that could have been used elsewhere; and (3) the value of lives of the projected fatalities. The benefit of a lockdown is calculated based on reducing the number of new infections going forward, and therefore avoiding a portion of these costs. Obviously, the longer the lockdown lasts, the larger the reduction in the number of new cases it will achieve. If a policymaker's only objective were to minimize the attack rate (the fraction of the population that will become symptomatically ill), the optimal solution would be to extend the lockdown until everyone is vaccinated. However, with each additional week of a lockdown the additional reduction in future infections becomes smaller, and since the benefits should be balanced against the costs to

[5] https://www.wsj.com/articles/long-a-holdout-from-covid-19-restrictions-sweden-ends-its-pandemic-experiment-11607261658.

Covid Economics 65, 20 January 2021: 78-107

the economy, a lockdown should be optimally stopped sooner. Using a range of reasonable assumptions, I find that a lockdown that starts a week from now is optimal because it produces a positive net benefit, and its optimal duration is between two and four weeks, depending on assumptions. I estimate that if no additional restrictions are imposed, even with the vaccination program currently in place, the pandemic will cost an additional \$2.4 trillion going forward if the value of statistical life (VSL) is used to value life and \$619 billion if life is valued with discounted quality-adjusted life years (dQALY).

Evidence shows that the lockdown measures adopted in parts of the United States and Europe in Spring 2020, which included bans on large social gatherings, closures of public places such as gyms, schools, bars and entertainment venues, and shelter-in-place orders, were highly successful at reducing the virus transmission rate (e.g., Courtemanche et al. (2020) and Flaxman et al. (2020)). I estimate that if the United States imposed a nation-wide lockdown similar to the lockdowns in Europe, which, depending on the assumptions, would optimally last between two and four weeks, it will generate a net benefit of up to \$1.2 trillion, or 6% of GDP.

II. The cost-benefit analysis

A. Estimating the future cost of the COVID pandemic

In order to estimate the dollar cost of the COVID-19 pandemic in the U.S., I follow the methodology used in studies of the costs of seasonal and hypothetical pandemic influenza outbreaks (e.g., Molinari et al. (2007) and CEA (2019)).[6]

[6]Throughout the paper, the terms "flu" and "influenza" are used interchangeably.

Covid Economics 65, 20 January 2021: 78-107

CEPR PRESS

A.1. Medical outcomes

An individual infected with the COVID-19 virus can have two outcomes: they can be asymptomatic or exhibit symptoms. Asymptomatic individuals do not miss work and do not incur any medical costs, although they can still infect others at the same rate as symptomatic individuals. Conditional on being symptomatic, an individual can have one of four progressively worse outcomes: (1) have mild symptoms and require no medical intervention, (2) have more severe symptoms and require an outpatient visit, (3) be hospitalized and survive, and (4) be hospitalized and die. Figure 1 plots the possible outcomes.

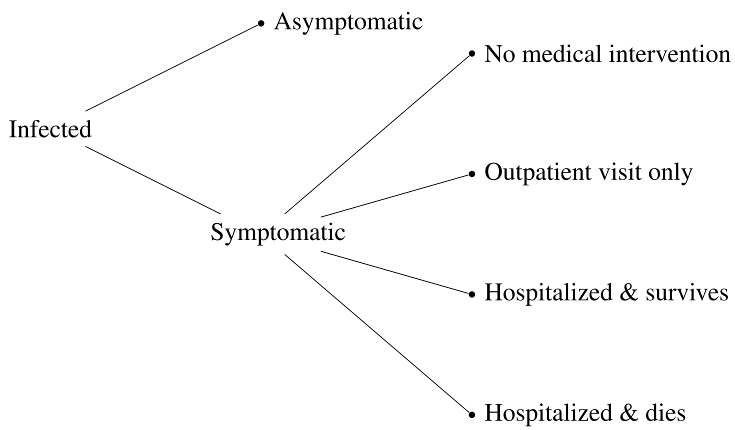


Figure 1. Outcomes for an infected person.
NOTES The figure presents possible outcomes for a person infected with the COVID-19 virus.

An important input into the analysis is the fraction of asymptomatic cases. Mizumoto et al. (2020) analyze the data from the quarantined Diamond Princess cruise ship and find that the asymptomatic fraction was 17.9%. However, given that the Diamond Princess sample consisted predominately of older adults, other studies have since estimated a higher fraction of asymptomatic infections among the general population. For example, Buitrago-Garcia et al. (2020) conduct meta-analysis of published papers using data around the world that they assess to be free of the sample

Covid Economics 65, 20 January 2021: 78-107

selection bias. They report a higher summary estimate of the proportion of the population that become infected with the virus and remain asymptomatic throughout the course of the infection of 31%. CDC's latest version of the "COVID-19 Pandemic Planning Scenarios" [7] also relies on meta-analysis of published papers to come up with an estimate for the asymptomatic fraction.[8] I use the assumptions from Scenario 5, "Current Best Estimate," that the asymptomatic fraction is 40%; it is derived as the mid-point of the estimates from published papers.[9]

Table I describes the probability that an infected person experiences each of the four possible outcomes of the disease as a function of their age and, when available, health risk status. Given that COVID risks increase with age, I divide the population into age groups (when an estimate for a particular age bin is not available from the literature, I calculate it by weighting the estimates for the overlapping age bins by the corresponding population fraction in each.) I obtain COVID hospitalization risks from Reese et al. (2020) and calculate the estimates for the older age bins using CDC's table on relative hospitalization risks by age.[10] I adjust these estimates for under-reporting by multiplying them by the ratio $2.5/7.7 = 0.32$ (Reese et al. (2020) find that hospitalized COVID cases are under-reported by a factor of 2.5 and overall COVID cases are under-reported by a factor of 7.7). Infection fatality rates (probability of dying conditional on being infected with COVID) are obtained from Levin et al. (2020), which is a meta-analysis of the literature and government reports that is restricted to studies of advanced economies, which includes only countries that currently belong in the Organization for Economic Cooperation and Development. In order to fill the more granular age bins in Table I, I also use the estimates from CDC's COVID-19 Pandemic Planning Scenarios, "Current Best Estimate."

Symptomatic individuals may fall into two groups: high- and low-risk. Patients who fall into high-risk health groups have pre-existing conditions that increase the likelihood of complications.

[7] https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html#five-scenarios.

[8] In doing so, CDC acknowledges the limitations of the current studies: "The percent of cases that are asymptomatic, i.e. never experience symptoms, remains uncertain. Longitudinal testing of individuals is required to accurately detect the absence of symptoms for the full period of infectiousness."

[9] This "best estimate" number aligns with estimates from Oran and Topol (2020), a meta-analysis that estimates the asymptomatic fraction to be 40% to 45%.

[10] Available from https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html.

Covid Economics 65, 20 January 2021: 78-107

The table provides cost estimates associated with each outcome for a symptomatic individual, as a function of age and health risk. Due to the lack of cost estimates specific to COVID infections and because COVID symptoms and the mode of transmission is similar to those of seasonal influenza,[11] I use the estimates for the proportion of high-risk individuals as well as medical and productivity costs from the seasonal influenza literature (Molinari et al. (2007) and CEA (2019)). However, early evidence indicates that COVID-19 may be more likely than influenza to leave survivors with long-term negative health effects,[12] which would cause me to underestimate the associated costs of an infection. Finally, for the calculation of the costs of lost productivity due to illness, I follow Barrot et al. (2020) and assume that a missed day of work represents productivity loss of $520.

For the individuals who die, society loses some productivity due to their inability to work during the period of the illness and, more importantly, the value of life. Policymakers employ several methods to estimate the value of life. Perhaps the most commonly used is the value of statistical life (VSL), which is estimated from studies assessing how much money people are willing to pay to increase the probability of staying alive. Following CEA (2019), I use inflation-adjusted VSL values by age group obtained from Aldy and Viscusi (2008), who estimate them from the wage premia paid by riskier jobs.[13] Because the value of medical costs are already factored into the value of life estimates, I do not add the medical costs for people who die. To calculate future costs of each pandemic management scenario considered, I calculate the total number of symptomatic individuals in each risk and age group that would fall into each of the four possible disease outcome categories and then multiply them by the associated costs, finally summing them up to obtain the total cost.

[11] See, e.g., the CDC description at `https://www.cdc.gov/flu/symptoms/flu-vs-covid19.htm`.

[12] See, e.g., `https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/late-sequelae.html`.

[13] While the authors are unable to estimate VSL for children, other studies obtain estimates from parents' willingness to pay for children's medical costs. The children's VSL estimate does not enter into the total cost calculation since COVID-19 studies assess a near-zero fatality risk for the younger age group.

Covid Economics 65, 20 January 2021: 78-107

A.2. The evolution of the pandemic curve

I use the SIR model to project the number of new COVID-19 infections at a weekly frequency. The starting point is 1/07/2021 (this is week 0), and I use the initial conditions as of this date to project the further evolution of the pandemic in the United States. I calculate the forward-looking costs from this time on and ignore the costs already incurred. Given the requirement for sick people to self-isolate for 14 days, I assume that a newly infected person is contagious for two weeks, during which time they will infect R_0 other people at the beginning of the pandemic, when no one in the population has immunity. (R_0 is called the basic reproduction number.) The number of other people that a contagious person infects is assumed to be spread evenly across the two weeks. Per SIR model, I assume that a recovered individual develops immunity and will not get infected or infect others, and the currently ill cannot be re-infected.

In addition to the immunity acquired by the recovered individuals, I account for the additional contribution to the population immunity from the ongoing vaccination program. Two COVID vaccines are already being administered, with more vaccine candidates going through the FDA approval process, and the stated goal to vaccinate 70% of the population[14] is expected to be achieved by June 2021.[15] I add the effect of vaccination to the SIR model by assuming that each person needs two vaccine doses spaced three weeks apart, at which point the vaccinated person is assumed to be fully immune and unable to spread the virus to others. I assume that vaccination starts with the most at-risk older population groups and progresses to younger groups,[16] with an equal number of people being vaccinated each week. I further model vaccinations as having started on December 14, when first doses of the Pfizer vaccine were administered and assume that vaccination will be completed by May 31, with 70% of the U.S. population fully vaccinated.

[14] https://www.commonwealthfund.org/publications/issue-briefs/2020/dec/how-prepared-are-states-vaccinate-public-covid-19.

[15] https://www.cnbc.com/2020/12/01/trump-covid-vaccine-chief-says-everyone-in-us-could-be-immunized-by-june.html.

[16] This is consistent with the CDC recommendations: https://www.cdc.gov/mmwr/volumes/69/wr/mm6949e1.htm.

Covid Economics 65, 20 January 2021: 78-107

Covid Economics 65, 20 January 2021: 78-107

To estimate the fraction of the population already recovered from COVID, it is important to account for under-reporting of COVID cases in the official statistics. I use the latest estimate of under-reporting from CDC, Reese et al. (2020), which analyzes four reasons for under-reporting—asymptomatic cases, symptomatic individuals not seeking medical attention, people seeking medical attention but not getting tested for COVID, and false negative test results—and estimates that due to these reasons, only one in 7.7 COVID cases ends up being detected and reported.[17] Given the overwhelming under-reporting of COVID cases, I assume that the vaccination program will not distinguish between the individuals who have not yet had COVID and those already recovered and immune.

A critical input into the SIR model is the virus R_0. CDC's "current best estimate" for the no-intervention COVID R_0 for the U.S. is 2.5.[18] However, increased sanitation, social distancing and the widespread use of face masks widely implemented in the United States were successful in reducing the virus reproduction number below this value. For example, Morley et al. (2020) study the effect of reduced personal mobility resulting form social distancing restrictions on the COVID reproduction number in several New York State counties. They use data from Unacast, a company that tracks and assigns letter grades to the reductions in mobility across various geographic areas; larger reductions in mobility are assigned higher grades. The figure presented on page 610 of the paper reports the effective reproduction rate that corresponds to each Unacast's mobility-reduction grade (I will use Panel B of the figure that removes outliers from the data). U.S.-wide mobility reduction roughly corresponds to Unacast's grade "D" assigned for a 40%–74% reduction in mobility. For example, Pishue (2020) finds that between March 14 and April 17 personal vehicle-miles traveled in the United States dropped by 46%, on average. Using a mobility index that aggregates cell phone data to capture changes in human movement over time, Archer et al. (2020) document a fluctuating but slightly larger drop in mobility, which is over 50% on average. Finally, Google's COVID-19 Community Mobility Report for the United States reports a similar-magnitude decline in the number of visits to public spaces. According to Morley et al. (2020), the "D" grade cor-

[17] As of January 7, 2021, CDC reports that there were 21.3 million COVID cases in the United States up to now, which implies that there were 163.7 million total infections after adjusting for under-reporting.

[18] `https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html#five-scenarios`.

responds to a reproduction number of roughly 1.75. Since the estimates in that paper were made during the early stages of the pandemic, when population-wide immunity was still low, I will use this number as my assumption for the basic reproduction number, R_0, that prevails with the interventions currently already in place. This R_0 estimate matches the currently observed data very well. Specifically, after inputting the number of recovered and therefore immune individuals into my SIR model with this R_0 parameter, I can match the number of people infected with COVID in the previous week and the current effective reproduction number, R_t (this variable measures the number of other people an infected person infects on average at time t in the pandemic, when a subset of the population has already recovered and gained immunity).[19]

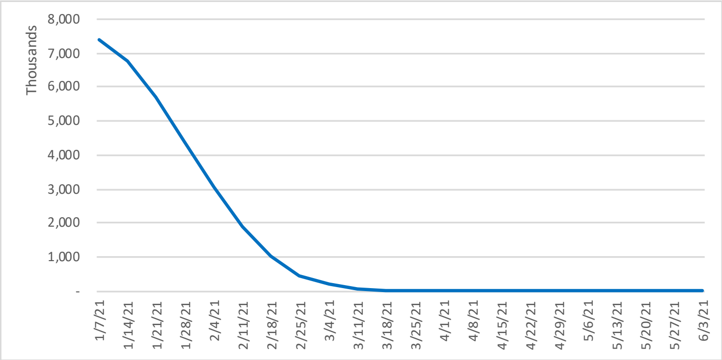


Figure 2. New symptomatic cases with no lockdown.
SOURCES CDC COVID Data Tracker, Census Bureau, author's calculations. SIR model inputs described in Table AI of the Appendix. **NOTES** The figure plots the predicted number of new symptomatic cases produced by SIR model. The model includes the effect of additional immunity acquired through vaccinations.

Figure 2 depicts a weekly-frequency projection of the number of new symptomatic COVID cases. It shows that with the current immune fraction of the population and the ongoing vaccination program, the number of new cases is projected to decline. Absent a lockdown, the pandemic is going to end roughly by mid-April 2021, and bring about 23.2 million additional symptomatic

[19] I estimate the current nation-wide R_t as the state-population-weighted average from the state-level median numbers reported on the website rt.live on January 7, 2021; the current R_t estimate is 1.04.

Covid Economics 65, 20 January 2021: 78-107

illnesses and 406 thousand additional deaths.[20] When the costs of symptomatic illnesses and the value of life are taken into account, these projections translate into a cost of \$2.42 trillion going forward, with medical costs and lost productivity contributing about 5% to this value and the rest attributed to the value-of-life losses.

B. Modeling the effect of a lockdown at different durations

By drastically reducing mobility, a lockdown holds the promise to significantly lower the virus reproductive number relative to the current value. It has been estimated that the Winter 2020 Wuhan lockdown reduced the COVID effective reproduction rate from above two to 0.3.[21] Given that lockdowns are likely to be less restrictive and more leniently enforced in Western countries, the drop in R_0 is likely to be more modest. Studies that analyze the effect of the Spring 2020 lockdowns in the United States and in Europe using mobility data, such as Google Community Mobility Index and smartphone GPS location data, find that lockdowns led to significant reductions in spatial movements (e.g., Pepe et al. (2020)). Lockdowns in Western countries were also found to generate large reductions in the virus reproduction rate. Using hospitalization records, Salje et al. (2020) estimate that in France the lockdown reduced the reproduction number by 77%, from 2.90 to 0.67. Flaxman et al. (2020) perform a broader analysis of the effect on the Spring 2020 lockdown across 11 European countries using data on COVID-related deaths and find that lockdowns on average decreased the virus reproduction rate by 81% to an average value of 0.66 across these countries. Using a survey, Jarvis et al. (2020) assess that in the U.K. the Spring 2020 lockdown lead to a 73% reduction in the number of contacts, which they estimate to reduce the R_0 from 2.6 prior to lockdown to a value of 0.62. Given the previously discussed estimate of the no-intervention $R_0 = 2.5$ for the United States, the average of these percent reductions implies that a lockdown will produce $R_0 = 0.58$. However, to err on the conservative side and because the U.K. is most culturally

[20] The projected fatality rate lower than the assumed average IFR because of the effect of the vaccination program that prioritizes the more vulnerable older population.

[21] https://qz.com/1834700/rt-the-real-time-r0-guiding-how-to-lift-coronavirus-lockdowns/.

Covid Economics 65, 20 January 2021: 78-107

similar to the United States, I will use the U.K. estimate and assume that a national lockdown in the U.S. will achieve $R_0 = 0.62$. In the sensitivity analyses, I use an even more conservative assumption that the basic reproduction number achievable with a lockdown is 25% higher, or 0.775.

I assume that if a lockdown were to be imposed, it would start a week from now, at the beginning of week 1. After a lockdown is lifted, the virus reproduction rate will revert to the pre-lockdown value. Figure 3 plots the incremental savings achieved from a lockdown as a function of the number of weeks that it is kept in place relative to the baseline of no lockdown depicted in Figure 2. The figure shows the savings increase with each additional week of a lockdown but at a declining rate.

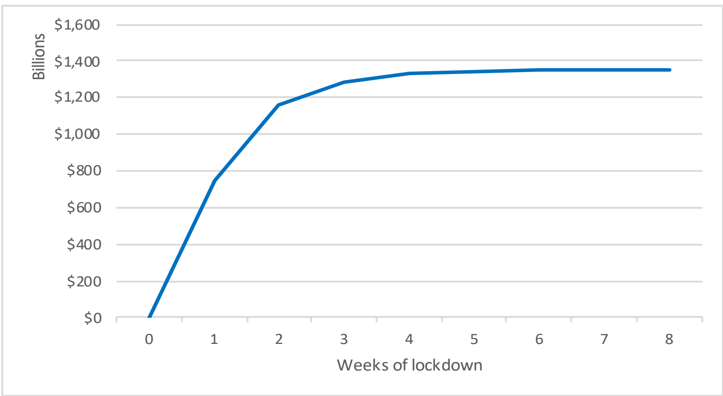


Figure 3. Savings from a lockdown as a function of its duration.
SOURCES CDC COVID Data Tracker, Census Bureau, author's calculations. SIR model inputs and other inputs used in the calculations are described in Table AI of the Appendix. **NOTES** The figure plots projected savings from a lockdown as a function of the number of weeks it is kept in place, assuming that it is imposed a week from now.

B.1. Incremental cost of a lockdown

When assessing the incremental impact of a lockdown on the economy, it is important to note that even in the absence of lockdown orders, a global pandemic depresses economic activity relative to normal times due to voluntary social distancing. Chen et al. (2020) collect a number of

Covid Economics 65, 20 January 2021: 78-107

Covid Economics 65, 20 January 2021: 78-107

high-frequency indicators of economic activity in the United States and Europe, such as electricity usage and mobility indicators, as well as additional economic indicators for the United States, such as unemployment insurance claims and employee-hours worked for small and medium sized businesses that employ hourly workers. The paper documents large reductions in mobility and economic activity even before the adoption of stay-at-home orders and nonessential business closures, and more so in places with more severe COVID outbreaks, indicating that people voluntarily limited their activities in order to protect themselves and others from the virus. Similarly, IMF (2020) uses high-frequency mobility indicators and shows that mobility decreases not only as a result of lockdown orders but also in response to rising COVID cases. The paper estimates that lockdown orders contributed about 40% and voluntary social distancing about 60% to the total decrease in mobility during lockdowns in advanced economies (Figure 2.2 of IMF (2020)).

I combine several estimates for the incremental cost that a lockdown would impose on the U.S. economy. Using a set of assumptions for which economic sectors would be affected and by how much, OECD (2020) estimates that for the G7 economies national lockdowns would cause annual GDP growth to decline by up to 2 percentage points per month of a lockdown. This translates into a GDP decline of \$107 billion per week for the United States ($0.5\% \times \$21.43$ trillion). However, this estimate is not incremental to the natural decline of economic activity caused by the pandemic. If, as discussed above, voluntary social distancing during the pandemic contributes about 60% to the reduction in economic activity, in line with the results discussed above, the incremental economic cost of an imposed lockdown will be about 40% of that estimate, or about \$43 billion per week.

Scherbina (2020) analyzes which sectors of the economy will be incrementally affected by a lockdown and by how much and also accounts for the additional costs of productivity losses caused by homeschooling demands on working parents. She estimates the incremental cost of a lockdown to be \$35.79 billion per week.

Barrot et al. (2020) obtain a slightly lower estimate for the lockdown cost, \$32 billion per week. Specifically, they estimate the number of workers in each U.S. state employed in the sectors that were closed in that state's Spring 2020 lockdown and who are unable to work from home (the total for the country is estimated to be 12.6 million workers) and multiply this number by the share of U.S. GDP per worker per week ($2,600) to arrive at the final estimate.

These three estimates are relatively close, and I will use the average of these estimates of $36.93 billion per week, when assessing the optimal lockdown duration. It must be noted that a number of non-economic costs and benefits of a lockdown have not been considered in the calculation above, as discussion in the Sensitivity Analyses subsection. Therefore, I also consider a more conservative assumption for the incremental cost of a lockdown, assuming that it is 25% higher than the estimate above.

B.2. Optimal lockdown duration

Optimally, the lockdown should end before its incremental benefit falls below its incremental cost to the economy. Figure 4 plots the incremental benefit of each additional week of a lockdown against its incremental cost to the economy. The incremental benefit line is declining, consistent with Figure 3 that shows that the incremental savings level off over time. The incremental savings line crosses the incremental cost line after four weeks. Therefore, four weeks is the optimal lockdown duration. After subtracting the incremental cost of the lockdown incurred over this time ($4 \times$ $36.93 bil.) from the incremental savings realized from preventing a subset of future infections, I estimate the associated net benefit relative to the baseline scenario of no lockdown to be $1.18 trillion, which is about 6% of GDP.

C. Sensitivity analyses

Given the uncertainty associated with some of my model inputs, I consider some alternative assumptions. Specifically, I consider changing the methodology for valuing life that would assign an even lower value to the lives of the elderly who are at a higher risk of COVID deaths, a lower assumption for lockdown effectiveness, a higher incremental cost, and a lower IFR.

Covid Economics 65, 20 January 2021: 78-107

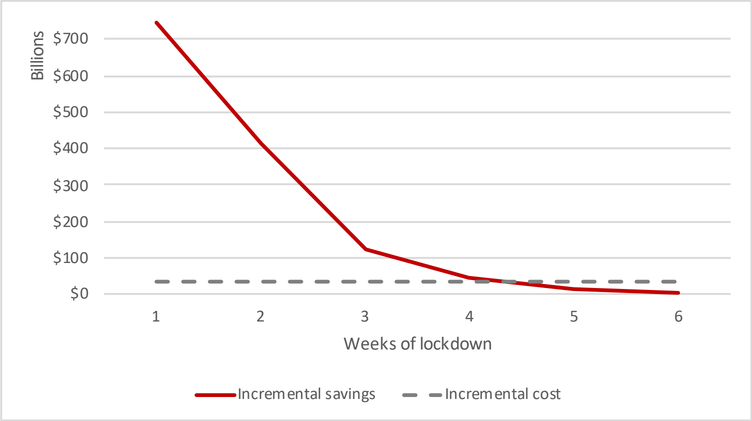


Figure 4. Incremental costs and savings of each additional week of a lockdown.
SOURCES CDC COVID Data Tracker, Census Bureau, author's calculations. Specific sources for SIR model inputs and incremental savings and cost estimates of a lockdown are provided in Table AI of the Appendix. **NOTES** Assuming that a lockdown is imposed a week from now, this figure plots projected incremental costs and savings of extending lockdown by another week, to the number of weeks specified on the horizontal axis.

C.1. Valuing life with discounted quality-adjusted life years

So far, I have used VSL to value life. However, health-adjusted life years (HALY) has been gaining popularity in recent years. Here I consider the quality-adjusted life years (QALY) methodology, which is perhaps the most widely used type of HALY (e.g., Neumann and Greenberg (2009)), Gold et al. (2002), Hubbell (2006) and Prieto and Sacristán (2003)). The morbidity or quality-of-life component is captured by a quality-of-life weight (QOL), which takes values between 0 and 1, with 0 representing death and 1 perfect health. The number of quality-adjusted life years (QALY) lived in one year is equal to the individual's QOL in that year. I follow Sassi (2006) to calculate the discounted value of all future quality-adjusted life years (dQALY):

$$\text{dQALY} = \sum_{t=a}^{a+L} \frac{QOL_t}{(1+r)^{t-a}}, \tag{1}$$

Covid Economics 65, 20 January 2021: 78-107

CEPR PRESS

where a is the current age, L is the residual life expectancy at age a, QOL_t is the expected health-related quality of life in year t, and r is the discount rate. I use a discount rate of 3% as is common in the literature (e.g., Sassi (2006) and Hubbell (2006)). The life expectancy for the U.S. population is obtained from Table VI of the 2020 National Vital Statistics Reports. I use cross-sectional average QOL weights by age estimated by Nyman et al. (2007) for the U.S. population using response data for the Medical Expenditure Panel Survey (Table 1 of the paper).

To translate the value of life into monetary terms when evaluating cost-effectiveness of medical interventions, the Institute for Clinical and Economic Review uses a range of $50,000-$150,000 per QALY. Song and Lee (2018) conduct a survey of the general Korean population and find that the willingness to pay for a cure treatment is more than twice as high as for a non-cure treatment. Since my objective is to value the lives of potential COVID fatalities, I will use the highest value of this range, that is $150,000 per QALY. The estimates of QALY-based monetary values of life are provided in Table II, estimated for the lowest bound of each age range.

Despite their increasing popularity, quality-adjusted life-year valuations (and HALY's more generally) have been criticized on technical and ethical grounds. By comparing dQALY values in Table II to VSL values in Table I, it can be seen that the former assign increasingly lower values to older age groups. The reason is that older people have fewer years of life remaining, and these years are of a worse quality due to deteriorating health. Likewise, dQALY would assign a lower value of life to people with disabilities and chronic health conditions relative to healthy people of the same age because the lower embedded QOL values.[22] On the technical side, a utility function has to have a very specific and, perhaps, unrealistic functional form, with features such as independence between life years and health status, in order to be consistent with the QALY maximization (e.g., Pliskin et al. (1980) and Prieto and Sacristán (2003)).

When the dQALY methodology is used to value life, absent a lockdown, the pandemic is projected to cost $619 billion going forward, with lost productivity and medical expenses representing 20% of this total and the value-of-life losses making up the rest. Table III presents the estimates

[22]See, e.g., Gold et al. (2002) for a discussion of the ethical challenges associated with using health-adjusted life expectancy to estimate the value of life.

Covid Economics 65, 20 January 2021: 78-107

of the optimal lockdown duration and the associated net savings by using dDALY instead of VSL to value life. It shows that a lockdown is still optimal under all assumptions considered, but its length is reduced by two weeks compared to when VSL is used. The net savings, computed as the incremental benefit achieved from reducing the number of future infections and deaths minus the incremental cost of the lockdown, are now substantially lower because of the lower value that this method assigns to the lives of older COVID victims.

C.2. Other incremental impacts of a lockdown

New literature has emerged that studies the non-economic effects of the Spring 2020 lockdowns. However, at this time, it may be too speculative to assign a dollar value to these additional effects that the literature finds since they are still imprecisely estimated.

Mental health. There is evidence that symptoms of depression and anxiety have increased during the lockdown. Pieh et al. (2020) evaluate several mental health and well-being indicators through an online survey with 1,006 respondents in the United Kingdom during the COVID-19 lockdown and find that the prevalence of depressive and anxiety symptoms increased relative to the pre-pandemic period. However, it is unclear from the study how much of this effect can be attributed to the pandemic itself and whether an incremental impact of a lockdown is positive or negative. Furthermore, in what could be interpreted as another negative indicator for mental health, the American Medical Association reports that the number of opioid and other drug-related deaths has increased during the Covid pandemic.[23] However, more research needs to be done to identify how much of the overdose increase can be attributed to the Spring lockdown.

Despite the evidence that lockdowns may have adverse effects on mental health, data does not show a positive association between lockdowns in suicides. Faust et al. (2020) study records from the Massachusetts Department of Health Registry of Vital Records and Statistics from January 2015 through May 2020 and find that suicide rates have actually decreased to 0.67 per 100,000

[23] https://www.ama-assn.org/system/files/2020-11/issue-brief-increases-in-opioid-related-overdose.pdf.

Covid Economics 65, 20 January 2021: 78-107

Covid Economics 65, 20 January 2021: 78-107

person-months from 0.81 per 100,000 person-months during the same period of 2019. Similarly, German data shows that suicides declined during the Spring 2020 lockdown relative to the same period of 2019, which may be partly explained by the lockdown's positive incremental effect on mental health, for example by reducing anxiety about being infected, eliminating commute to work, and allowing more time with family.[24] More research is needed to more precisely isolate the incremental effect of lockdowns on mental health.

Traffic injuries and fatalities. A clear benefit of a lockdown is a reduction in traffic and the ensuing decline in the number of traffic injuries and fatalities, which were estimated to fall by half in California during the Spring 2020 lockdown (Shilling and Waetjen (2020)). For the country as a whole, the National Highway Traffic Safety Administration estimates that for the April–June 2020 period, 302 fewer traffic deaths were recorded relative to the same period in 2019.[25] This reduction in traffic deaths translates into a VSL benefit of $202 million for each week of the lockdown (using the population-average VSL of $8.68 million), and the benefit is even higher if the medical costs of caring for the injured are taken into account.

Environmental impact. Another benefit of reduced traffic is less pollution, which has health benefits. Venter et al. (2020) utilize a network of air quality stations distributed across 34 countries to measure the air quality during the lockdown period up until May 15, 2020. They estimate that the air quality has substantially improved, which resulted in large public health benefits. The study estimates that a total of 49,900 pollutant-related deaths and 89,000 pediatric asthma emergency room visits were avoided in the 34 countries in the study sample. Similarly, Archer et al. (2020) document that the reduction in personal mobility caused by the COVID pandemic lead to a significant decrease in NO_2 concentrations in the United States.

Crime. Spring lockdowns are documented to have reduced the overall crime rate but increased the incidence of domestic violence. Citing the "opportunity theory" of crime, which posits that restrictions on mobility and social interactions will present fewer opportunities for criminal activity,

[24] https://www.dw.com/en/is-social-distancing-during-coronavirus-causing-more-suicides/a-53584282.

[25] https://www.reuters.com/article/us-health-coronavirus-usa-traffic-exclus/u-s-traffic-deaths-fell-after-coronavirus-lockdown-but-drivers-got-riskier-idUSKBN26M6KR.

Covid Economics 65, 20 January 2021: 78-107

the United Nations Office on Drugs and Crime (UNODC (2020)) investigates crime rates around the world and documents that reported incidents of robbery, theft and burglary declined by more than 50 per cent in most countries, and the declines were larger in countries with stricter lockdown regimes (however, the report notes that some of the decline may be attributed to under-reporting). The number of homicides has also greatly declined, but only in some countries, and started rebounding once lockdown measures were relaxed.

These findings are largely corroborated in Bullinger et al. (2020), who analyze crime data and 911 calls made in Chicago during the Spring 2020 stay-at-home orders. They find that overall crime-related arrests decreased by 57%[26] and 911 calls decreased by 6% during this period. However, 911 calls reporting domestic violence have increased by 7%, though domestic-violence-related arrests decreased by 27%. The authors speculate that the decrease in domestic violence arrests may be partly explained by under-reporting. The increase in the number of 911 calls reporting domestic violence during the pandemic is consistent with evidence presented in other papers (e.g., Leslie and Wilson (2020)).

Reduction in overall mortality. Kung et al. (2020) show that the New Zealand lockdown lead to an 11% decrease in the weekly death rate relative to historical trends. The authors provide evidence that this reduction was largely explained by the reduction in seasonal influenza and pneumonia, though other factors, such as fewer traffic deaths, reduced air pollution and lower occupational hazards, likely played a role as well.

Additional incremental effects. A lockdown would likely have additional shorter- and longer-term costs and benefits. One clear benefit is that an increased reliance on technology will help boost future GDP growth. An increased ability to work remotely will allow more individuals to enter the workforce, and companies will be able achieve a higher return on investment by saving on real estate leases and travel costs. Moreover, with less commuting and reduced traffic, employees may gain productive hours.

[26]This decrease may be partly explained by a new policy to limit or halt prosecutions of low-level, non-violent offenses, adopted by the Chicago Police Department on March 20 in an effort to protect first responders.

CEPR PRESS

Lockdowns also have a number of negative incremental effects, in addition to the ones already mentioned. Reduced access to medical services may lead to negative health consequences in the longer term. A lower education quality may result in a marginally less productive future workforce. Lockdowns may cause an incremental increase in the number of bankruptcies, resulting in deadweight losses associated with a less efficient re-deployment of business assets. In the sensitivity analyses presented in Table III, I make a more conservative assumption for the incremental cost of a lockdown by increasing it by 25%, to $46.16 billion a week. The table shows that for the main results, the higher incremental cost does not shorten the optimal lockdown duration but reduces the estimated net savings.

Covid Economics 65, 20 January 2021: 78-107

C.3. Reducing the IFR

The IFR estimates in the paper are based on meta-analyses of academic papers and government data for developed economies. The implied population-weighted average IFR of 1.33% is consistent with a number of other estimates for the COVID mortality rate. However, I also try a more conservative estimate, by assuming that the IFR for each age bin is reduced by 25%, thus resulting in a population-weighted average IFR of 1.00%. Under this assumption, absent a lockdown, the future death toll is projected to be almost 305 thousand, and the future cost of the pandemic to be $1.84 trillion, with medical costs and lost productivity representing 6% of this number and the rest attributed to value-of-life losses. When dQALY are used to value life, the pandemic is projected to cost $490.39 billion going forward, with medical and productivity costs representing 24% of this number.

Table AII in the Appendix reports the estimates of the optimal lockdown duration and the corresponding net savings for this IFR assumption. Compared to the main results, the optimal lockdown duration is shortened by one to two weeks when VSL is used to value life and by one week for two sets of assumptions when dQALY is used. And since projected fatalities are lower, the estimated net benefits of a lockdown are reduced.

D. Limitations

I use a number of parameters reported in the COVID literature as inputs into my analysis, such as the magnitude of under-reporting of cases, the fraction of asymptomatic cases, the hospitalization rate, and the infection fatality ratio (IFR). Therefore, I would like to caveat the findings by pointing out that these inputs may be imprecisely estimated. For example, robustness analyses show that the optimal lockdown duration may be shorter if the IFR is lower. If the speed of vaccinations is slower than what I assume, or if vaccinations do not prioritize the older population, the optimal lockdown duration may be longer. Additionally, early evidence is emerging on the late sequelae of COVID-19.[27] If the long-term consequences of COVID are severe enough to result in large medical expenses, productivity losses, and shortened life spans, the optimal lockdown duration should be longer than estimated in order to help further minimize the number of new infections.[28]

Another critical parameter is the reduction in the virus transmission that can be achieved with a national lockdown. I rely on estimates obtained from studies of the Spring 2020 European lockdowns, which may not be perfectly applicable to the United States. In the sensitivity analyses, I consider a more conservative assumption.

The incremental costs that a lockdown would impose on the economy may be imprecisely estimated. In the sensitivity analyses, I use a more conservative cost estimate and still find that a lockdown would be beneficial, albeit with a shorter optimal length in some specifications. It is also possible that the incremental costs to the economy may be increasing with each subsequent week of a lockdown, perhaps through the higher likelihood of bankruptcies and the associated dead-weight losses. In that case, the optimal lockdown duration may be shorter than estimated.

New literature has emerged on the non-economic costs and benefits of a lockdown, such as its impact on crime, air pollution, mental health, etc. Overall, this literature finds a number of positive or ambiguous incremental effects. Presently, due to the lack of precise estimates, it is unclear

[27] See, e.g., `https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/late-sequelae.html`.

[28] Cutler and Summers (2020) estimate that long-term health impairments may reduce quality-adjusted life expectancy of COVID survivors by 35%.

Covid Economics 65, 20 January 2021: 78-107

how to assign a monetary value to these incremental effects. However, if those were taken into consideration, they may help reduce the estimated cost of a lockdown, which is calculated purely as the cost to economic activity.

Finally, I do not model the worsening of medical outcomes as a function of the number of new infections. A high number of new infections may overwhelm the medical system and result in a higher likelihood for severe medical outcomes. Taking this effect into account would make a lockdown more beneficial and potentially extend its optimal duration.

III. Conclusion

To my knowledge, this paper presents the first attempt to examine whether it is beneficial for the United States to follow the lead of a number of European countries and order a national lockdown and to estimate its optimal duration while using current conditions and explicitly modeling the ongoing vaccination program. I find that even with ongoing vaccinations a lockdown will generate significant net benefits and that it should optimally last up to four weeks. When I use a more conservative approach to valuing lives, using discounted quality-adjusted life-years that assigns a significantly lower values to the lives of older individuals, I find that a lockdown is still beneficial but that its optimal duration decreases to two weeks. Additionally, when I consider more conservative assumptions for the lockdown effectiveness, its incremental cost to the economy, and IFR, I still find that a lockdown would be optimal, albeit at a shorter duration.

A number of additional arguments can be made in favor a lockdown. First, the vaccination program is currently progressing at a substantially slower speed than initially promised. With the vaccination end-goal shifted further into the future, additional non-pharmaceutical interventions would be even more helpful. Second, adding to the costs of the pandemic is the emerging evidence of serious long-term complications resulting from COVID infections. If those lead to large productivity losses and medical expenses and shortened life spans, the associated costs should be added to the estimated cost of the pandemic. Third, reports are currently emerging that hospitals

Covid Economics 65, 20 January 2021: 78-107

are preparing to ration care due to an influx of COVID cases.[29] While currently outside of the model, a lockdown will help reduce the pressure on the medical system and result in better medical outcomes for COVID patients. Fourth, the literature on the Spring 2020 lockdowns documents a number of positive effects for public health and well-being, such as reduced traffic accidents, lower pollution, and lower mortality due to influenza and pneumonia. While some effects, such as domestic violence and mental health are negative, the overall effect is likely positive. Taken together, these arguments imply higher lockdown benefits and lower incremental costs than those used in the paper and point to a longer optimal lockdown duration.

Despite the obvious benefits, there is widespread reluctance to impose additional mobility restrictions in the United States. COVID presents a low threat to the young and healthy but a high threat to the elderly and those with underlying medical conditions. And while a lockdown may not benefit each individual it will benefit society as a whole, as the analysis in this paper shows.

[29] https://www.healthline.com/health-news/hospitals-may-have-to-ration-care-as-covid-19-hits-record-highs.

Covid Economics 65, 20 January 2021: 78-107

References

Aldy, J. and W. Viscusi (2008). Adjusting the value of a statistical life for age and cohort effects. *Review of Economics and Statistics 90*(3), 573–581.

Archer, C. L., G. Cervone, M. Golbazi, N. A. Fahel, and C. Hultquist (2020). Changes in air quality and human mobility in the usa during the covid-19 pandemic. *Bulletin of Atmospheric Science and Technology*.

Arias, E. and J. Xu (2020). *National Vital Statistics Reports. United States Life Tables, 2018*.

Barrot, J.-N., B. Grassi, and J. Sauvagnat (2020). Costs and benefits of closing businesses in a pandemic. Working paper.

Boehmer, T. K., J. DeVies, E. Caruso, K. L. van Santen, S. Tang, C. L. Black, K. P. Hartnett, A. Kite-Powell, S. Dietz, M. Lozier, and A. V. Gundlapalli (2020). Changing age distribution of the covid-19 pandemic — united states, may–august 2020. *Morbidity and Mortality Weekly Report, US Department of Health and Human Services/Centers for Disease Control and Prevention 69*(39).

Buitrago-Garcia, D., D. Egli-Gany, M. J. Counotte, S. Hossmann, H. Imeri, A. M. Ipekci, G. Salanti, and N. Low (2020). Occurrence and transmission potential of asymptomatic and presymptomatic sars-cov-2 infections: A living systematic review and meta-analysis. *PLoS Medicine 17*(9), 1–25.

Bullinger, L. R., J. B. Carr, and A. Packham (2020). Covid-19 and crime: Effects of stay-at-home orders on domestic violence. NBER Working paper.

CEA (2019). *Mitigating the Impact of Pandemic Influenza through Vaccine Innovation*, `https://www.whitehouse.gov/wp-content/uploads/2019/09/Mitigating-the-Impact-of-Pandemic-Influenza-through-Vaccine-Innovation.pdf`. Council of Economic Advisers, White House.

Chen, S., D. Igan, N. Pierri, and A. F. Presbitero (2020). Tracking the economic impact of covid-19 and mitigation policies in europe and the united states. *Covid Economics 36*, 1–24.

Courtemanche, C., J. Garuccio, A. Le, J. Pinkston, and A. Yelowitz (2020). Strong social distancing measures in the united states reduced the covid-19 growth rate. *Health Affairs 39*(7), 1237–1246.

Cutler, D. M. and L. H. Summers (2020). The covid-19 pandemic and the $16 trillion virus. *The Journal of the American Medical Association 324*(15), 1495—1496.

Faust, J. S., S. B. Shah, C. Du, S.-X. Li, Z. Lin, and H. M. Krumholz (2020). Suicide deaths during the stay-at-home advisory in massachusetts. *medRxiv*.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, I. C. C.-. R. Team, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, and S. Bhatt (2020). Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature 584*, 257–261.

Gallup, G. and F. Newport (1991). Mirror of america: Fear of dying. *Gallup Poll News Services*.

Gold, M. R., D. Stevenson, and D. G. Fryback (2002). Halys and qalys and dalys, oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health 23*(1), 115–134. PMID: 11910057.

Covid Economics 65, 20 January 2021: 78-107

CEPR PRESS

Hubbell, B. J. (2006). Implementing qalys in the analysis of air pollution regulations. *Environmental & Resource Economics 34*, 365–384.

IMF (2020). *World Economic Outlook, Chapter 2, "The Great Lockdown: Dissecting the Economic Effects," October 2020.* International Monetary Fund.

Jarvis, C. I., K. van Zandvoort, A. Gimma, K. Prem, C. nCov working group, P. Klepac, G. J. Rubin, and W. J. Edmunds (2020). Quantifying the impact of physical distance measures on the transmission of covid-19 in the uk. *medRxiv*.

Kung, S., M. Doppen, M. Black, T. Hills, and N. Kearns (2020). Assessing the age specificity of infection fatality rates for covid-19: systematic review, meta-analysis, and public policy implications. *Lancet*.

Leslie, E. and R. Wilson (2020). Sheltering in place and domestic violence: Evidence from calls for service during covid-19. *Journal of Public Economics*. Forthcoming.

Levin, A. T., W. P. Hanage, N. Owusu-Boaitey, K. B. Cochran, S. P. Walsh, and G. Meyerowitz-Katz (2020). Assessing the age specificity of infection fatality rates for covid-19: systematic review, meta-analysis, and public policy implications. *European Journal of Epidemiology*.

Mizumoto, K., K. Kagaya, A. Zarebski, and G. Chowell (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, available at https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180. *Euro Surveill 10*(25).

Molinari, N., I. Ortega-Sanchez, M. Messonnier, W. Thompson, P. Wortley, and et al. (2007). The annual impact of seasonal influenza in the u.s.: Measuring disease burden and costs. *Vaccine 25*(27), 5086–96.

Morley, C. P., K. B. Anderson, J. Shaw, T. Stewart, S. J. Thomas, and D. Wang (2020). Social distancing metrics and estimates of sars-cov-2 transmission rates: Associations between mobile telephone data tracking and r. *Journal of Public Health Management and Practice 26*(6), 606–612.

Neumann, P. J. and D. Greenberg (2009). Is the united states ready for qalys? *Health Affairs 24*(5), 1366–1371.

Nyman, J. A., N. A. Barleen, B. E. Dowd, D. W. Russell, S. J. Coons, and P. W. Sullivan (2007, July). Quality-of-life weights for the us population: self-reported health status and priority health conditions, by demographic characteristics. *Medical Care 45*, 618–628.

OECD (2020). *Evaluating the Initial Impact of COVID Containment Measures on Activity, June 10, 2020,* *https://read.oecd-ilibrary.org/view/?ref=126_126496-evgsi2gmqj&title=Evaluating_the_initial_impact_of_COVID-19_containment_measures_on_economic_activity*. Organization for Economic Co-operation and Development.

Oran, D. P. and E. J. Topol (2020). Prevalence of asymptomatic sars-cov-2 infection : A narrative review. *Annals of Internal Medicine 173*(5), 362–367.

Pepe, E., P. Bajardi, L. Gauvin, F. Privitera, B. Lake, C. Cattuto, and M. Tizzoni (2020). Covid-19 outbreak response, a dataset to assess mobility changes in italy following national lockdown. *Scientific Data 7*(230), 1–7.

Covid Economics 65, 20 January 2021: 78-107

Pieh, C., S. Budimir, J. Delgadillo, M. Barkham, J. R. J. Fontaine, and T. Probst (2020, October). Mental health during covid-19 lockdown in the united kingdom. *Psychosomatic Medicine*.

Pishue, B. (2020). COVID-19's impact on freight: an analysis of long-haul freight movement during a pandemic. Technical report, INRIX Research Report.

Pliskin, J. S., D. S. Shepard, and M. C. Weinstein (1980). Utility functions for life years and health status. *Operations Research 28*(1), 206–224.

Prieto, L. and J. A. Sacristán (2003). Problems and solutions in calculating quality-adjusted life years (qalys). *Health and Quality of Life Outcomes 1*(80), 1366–1371.

Reese, H., A. D. Iuliano, N. N. Patel, S. Garg, L. Kim, B. J. Silk, A. J. Hall, A. Fry, and C. Reed (2020). Estimated incidence of covid-19 illness and hospitalization — united states. *Clinical Infectious Diseases*.

Salje, H., C. T. Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hozé, J. Richet, C.-L. Dubost, Y. L. Strat, J. Lessler, D. Levy-Bruhl, A. Fontanet, L. Opatowski, P.-Y. Boelle, and S. Cauchemez (2020). Estimating the burden of sars-cov-2 in france. *Science 369*(6500), 208–211.

Sassi, F. (2006). Calculating qalys, comparing qaly and daly calculations. *Health Policy and Planning 21*, 402–408.

Scherbina, A. (2020). Determining the optimal duration of the covid-19 suppression policy: A cost-benefit analysis. Working paper.

Shilling, F. and D. Waetjen (2020). *Impact of COVID19 Mitigation on California Traffic Crashes,* `https://roadecology.ucdavis.edu/files/content/projects/COVID_CHIPs_Impacts_updated_415_0.pdf`. UC Davis.

Song, H. J. and E.-K. Lee (2018). Evaluation of willingness to pay per quality-adjusted life year for a cure. a contingent valuation method using a scenario-based survey. *Medicine 97*(38), 1–10.

Tumpey, A. J., D. Daigle, and G. Nowak (2018). *Communicating During an Outbreak or Public Health Investigation,* `https://www.cdc.gov/eis/field-epi-manual/chapters/Communicating-Investigation.html`. Centers for Disease Control and Prevention.

UNODC (2020). *Effect of the COVID-19 pandemic and related restrictions on homicide and property crime*. United Nations Office on Drugs and Crime.

Venter, Z. S., K. Aunan, S. Chowdhury, and J. Lelieveld (2020, November). Air pollution declines during covid-19 lockdowns mitigate the global health burden. *Environ Research*.

Covid Economics 65, 20 January 2021: 78-107

Covid Economics 65, 20 January 2021: 78-107

Table I
Risks and Costs Associated with a COVID Infection

	Age Group						
	0-19	20-44	45-54	55-64	65-74	75-84	≥85
% of US population	25%	33%	13%	13%	9%	5%	2%
OUTCOME PROBABILITIES							
Clinical outcomes							
Prob. of hospitalization	1.29%	2.27%	4.23%	6.17%	13.96%	22.34%	60.00%
Probability of dying (IFR)	0.003%	0.020%	0.423%	0.500%	2.500%	8.500%	28.300%
Outcomes for symptomatic patients							
Proportion high-risk	8%	15%	24%	33%	51%	51%	51%
Outpatient visit							
- low-risk patients	32%	32%	32%	31%	62%	62%	62%
- high-risk patients	77%	63%	63%	63%	82%	82%	82%
COST ESTIMATES							
Case not medically attended							
Medical cost (all risk)	$5	$5	$5	$5	$5	$5	$5
Lost productivity (all risk)	$260	$260	$260	$260	$520	$520	$520
Outpatient visit							
Low-risk medical cost	$161	$212	$233	$254	$410	$410	$410
Low-risk lost productivity	$520	$520	$780	$1,040	$1,560	$1,560	$1,560
High-risk medical cost	$1,098	$1,227	$1,234	$1,240	$806	$806	$806
High-risk lost productivity	$2,080	$1,040	$1,560	$2,080	$3,640	$3,640	$3,640
Hospitalization							
Low-risk medical cost	$25,408	$32,174	$34,960	$37,745	$19,379	$19,379	$19,379
Low-risk lost productivity	$4,680	$6,240	$6,500	$6,760	$6,760	$6,760	$6,760
High-risk medical cost	$70,938	$80,760	$75,334	$69,908	$28,346	$28,346	$28,346
High-risk lost productivity	$11,960	$10,920	$11,700	$12,480	$9,360	$9,360	$9,360
Fatalities							
Low-risk lost productivity	$4,680	$6,240	$6,500	$6,760	$6,760	$6,760	$6,760
High-risk lost productivity	$11,960	$10,920	$11,700	$12,480	$9,360	$9,360	$9,360
Value of statistical life ($, mil.)	5.76	12.34	10.05	7.75	5.29	5.29	5.29

SOURCES For hospitalization risks: Reese et al. (2020) and CDC; for IFR: Levin et al. (2020) and CDC's "COVID-19 Pandemic Planning Scenarios," Scenario 5; for high- and low-risk probabilities and cost estimates: Molinari et al. (2007) and CEA (2019); for productivity losses: Molinari et al. (2007) and Barrot et al. (2020); for VSL: Aldy and Viscusi (2008) and CEA (2019); author's calculations. **NOTES** This table presents the risks and per-person medical risks and productivity costs associated with various outcomes of the COVID-19 infection.

Covid Economics 65, 20 January 2021: 78-107

Table II
Value of Discounted Quality-Adjusted Life Years, by Age

	Age Group						
	0-19	20-44	45-54	55-64	65-74	75-84	≥85
Value of dQALY (($, mil.)	4.05	3.44	2.73	2.22	1.71	1.14	0.82

SOURCES Nyman et al. (2007) for QOL weights; National Vital Statistics Reports for life expectancy by age; author's calculations. **NOTES** This table presents the dollar value of discounted quality-adjusted life years, calculated at the lower boundary of each age group. The discount rate is 3% per year, and the dollar value of QALY is $150,000.

Table III
Sensitivity of the optimal lockdown duration to alternative assumptions

Assumptions			Optimal	Net savings relative to
Lockdown R_0	Life is valued with	Incremental cost of lockdown	lockdown duration	no-lockdown baseline ($, billion)
0.620	VSL	$36.93 bil.	4 weeks	$1,184.89
		$46.16 bil.	4 weeks	$1,147.96
	dQALY	$36.93 bil.	2 weeks	$220.56
		$46.16 bil.	2 weeks	$202.09
0.775	VSL	$36.93 bil.	4 weeks	$1,051.19
		$46.16 bil.	4 weeks	$1,014.26
	dQALY	$36.93 bil.	2 weeks	$182.31
		$46.16 bil.	2 weeks	$163.85

SOURCES Author's calculations. **NOTES** This table presents the optimal lockdown duration as a function of the assumptions listed in the table. The right-hand column presents the incremental net savings of a lockdown of optimal length calculated as its benefits minus the associated costs incurred over the lockdown duration.

CEPR PRESS

Covid Economics 65, 20 January 2021: 78-107

Appendix for
"Could the United States benefit from a lockdown? A cost-benefit analysis"

Table AI
Summary of variables, sources, and assumptions used in the paper

COVID basic reproduction number with social distancing measures currently in place, R_0 (1.75). Sources: Morley et al. (2020) for the COVID reproduction number as a function of reduced mobility and Archer et al. (2020), Pishue (2020), and Google's COVID-19 Community Mobility Report for mobility reduction estimates for the United States.

COVID effective reproduction number, R_t (1.04). Source: website rt.live and author's calculations.

COVID basic reproduction number with a national lockdown (0.62). Source: Jarvis et al. (2020). A 25% higher estimate, 0.775, is used in sensitivity analyses.

COVID case under-reporting (true number of COVID cases is 7.7 times higher than official statistics). Source: Reese et al. (2020).

Number of COVID cases. Source: CDC COVID Data Tracker, adjusted for under-reporting.

Vaccination. Assume that vaccinations started on 12/14/2020 and will be completed by 05/31/2020, with the objective to get 70% of the population fully vaccinated by that date. Vaccination involves two vaccine doses, administered three weeks apart. The vaccinated cannot get infected with or spread the virus to others. Vaccination is assumed to progress at a constant speed, with the same number of people immunized each week, from oldest to youngest population groups. Due to the inability to determine who has already recovered from COVID, vaccinations include recovered individuals as well.

Fraction of asymptomatic cases (40%). Source: CDC "COVID-19 Pandemic Planning Scenarios," Scenario 5: "Current Best Estimate."

Medical costs and outcomes (as reported in Table I). Sources: For hospitalization risks: Reese et al. (2020), Table 1, with missing age bins augmented by CDC's estimates on relative hospitalization risks by age (`https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html`), with all estimates multiplied by 0.32 to adjust for hospital and overall case under-reporting per Reese et al. (2020); for IFR: Levin et al. (2020) and the CDC "COVID-19 Pandemic Planning Scenarios," Scenario 5: "Current Best Estimate." A 25% lower IFR estimate for each age bin is used in sensitivity analyses. For high- and low-risk probabilities and cost estimates: Molinari et al. (2007) and CEA (2019), Table 2.

CEPR PRESS

VSL (as reported in Table I). Sources: Aldy and Viscusi (2008) and CEA (2019).

dQALY (as reported in Table II). Sources: Nyman et al. (2007) for QOL weights, National Vital Statistics Reports for life expectancy by age, 3% discount rate, $150,000 value per QALY, author's calculations.

Incremental cost of a lockdown ($36.93 billion per week). Source: average estimate from OECD (2020), Scherbina (2020), and Barrot et al. (2020). A 25% higher estimate, $46.16 billion per week, is used in sensitivity analyses.

Covid Economics 65, 20 January 2021: 78-107

Table AII
Sensitivity of the optimal lockdown duration to alternative assumptions, with IFR reduced by 50%

Assumptions			Optimal lockdown duration	Net savings relative to no-lockdown baseline ($, billion)
Lockdown R_0	Life is valued with	Incremental cost of lockdown		
0.620	VSL	$36.93 bil.	3 weeks	$866.76
		$46.16 bil.	2 weeks	$791.13
	dQALY	$36.93 bil.	2 weeks	$159.31
		$46.16 bil.	1 week	$104.04
0.775	VSL	$36.93 bil.	3 weeks	$759.12
		$46.16 bil.	2 weeks	$676.47
	dQALY	$36.93 bil.	2 weeks	$129.02
		$46.16 bil.	1 week	$82.51

SOURCES Author's calculations. **NOTES** This table is an update of Table III in the main text based on the assumption that IFR for each age bin is reduced by 25%. It presents the optimal lockdown duration as a function of assumptions listed in the table. The right-hand column presents the incremental net savings of a lockdown of optimal length calculated as its benefits minus the associated costs incurred over the lockdown duration.

CEPR PRESS

Covid Economics **Issue 65, 20 January 2021**

Information processing skills of short sellers: Empirical evidence from the Covid-19 pandemic[1]

Levy Schattmann,[2] Jan-Oliver Strych[3] and Joakim Westerholm[4]

Date submitted: 9 January 2021; Date accepted: 12 January 2021

We aim to answer if superior performance by short sellers' is generated by processing public information rather than by exploiting private information. To achieve this, we analyze if short sellers with healthcare expertise outperform in short selling of non-healthcare stocks compared to those with no healthcare expertise. Since we expect that any short sellers' private information about healthcare stocks is unlikely to be material for non-healthcare stocks, we conclude that any observed outperformance in non-healthcare stocks is more likely caused by processing public information. As an identification strategy, we interpret the outbreak of the Covid-19 pandemic as a treatment to short sellers with healthcare expertise. Our measures of healthcare expertise are based on pre-Covid-19 performance related to either holding or covering a short position in healthcare stocks. Using a unique German sample of daily short selling data, we find that treated short positions identified by general shorting (covering) out-performance are associated with lower 10-day CARs for non-healthcare stocks by an economi-cally significant magnitude of 4.3 percent (7.2 percent). Robustness test rule out that our results are also driven by the use of private information or non information-based trading advantages such as better funding or lending ability of observed short sellers.

1 We thank Jean-Gabriel Cousin, Gael Imad'Eddine, and Martin Ruckes, and participants of the Research Semi-nar 2020 at the University of Lille for helpful comments and discussions.

2 Graduate student researcher, Department for Finance and Banking, Karlsruhe Institute of Technology.
3 Assistant Professor of Finance, Department for Finance and Banking, Karlsruhe Institute of Technology.
4 Associate Professor of Finance, Sydney University Business School.

Covid Economics 65, 20 January 2021: 108-152

Covid Economics 65, 20 January 2021: 108-152

1. Introduction

The literature provides overwhelming evidence that short sellers are informed and sophisticated traders with an information advantage over other market participants (e.g., Asquith, Pathak, and Ritter (2005); Boehmer, Jones, Wu and Zhang (2020); Boehmer, Jones, and Zhang (2008); Desai, Ramesh, Thiagarajan, and Balachandran (2002); Diether, Lee, and Werner (2009)). But how do they obtain their information advantage? Whereas there are some empirical indications for the use of private information for informed short selling (e.g., Boehmer, Jones, Wu, and Zhang (2020); Christophe, Ferri, and Angel (2004); Karpoff and Lou (2010)), Engelberg, Reed, and Ringgenberg (2012), for instance, document that short seller's information advantage is determined by their superior skills to process public information.

Our paper aims to provide empirical evidence for this information processing argument and applies a difference-in-differences approach by interpreting the exogenous shock event of the Covid-19 pandemic in 2020 as treatment to short sellers with healthcare expertise. If short sellers with healthcare expertise (treatment group) outperform a control group of other short sellers without healthcare-specific trading skills in short selling non-healthcare stocks after the outbreak of the Covid-19 pandemic, we interpret this finding as empirical evidence that this outperformance is likely caused by public information processing skills of treated short sellers rather than their use of private information on healthcare stocks because private information on healthcare stocks is less likely applicable in superior trading of non-healthcare stocks. As outcome of superior information processing skills we regard better predictions of the pandemic's impact on non-healthcare firms' stock performance because of two reasons. First, we expect that healthcare expertise leads to more timely acquisition of publicly-available but hard-to-find pandemic-related information. Second, we assume that healthcare expertise enables a more accurate understanding and subsequent prediction of the dissemination and health impact of Covid-19 to anticipate more precisely customer behavior changes and governmental measures such as lockdowns, shutdowns, stay-at-home orders, and travel restrictions.

We use the Covid-19 pandemic in our analysis as an appropriate identification strategy of public information processing skills as source of short sellers' information advantage for several reasons. First, it is exogenous by nature so no short sellers have anticipated it. More precisely, healthcare expertise is unlikely to enable them to anticipate it so the group assignment is uncorrelated with this treatment event. Second, the newness of the Covid-19 disease makes it rather unlikely that private information on healthcare stocks helps anticipate the availability of a vaccine or any medical treatment against Covid-19 shortly after its outbreak so that we assume that outperformance of short sellers with healthcare expertise is not driven by such private information. Third, since we rely on timely disclosed public information in our study, Covid-19 caused skyrocketing volatility and higher short selling constraints through stock recalls by selling long investors and withdrawals in short sellers' funds likely prevent uninformed short sellers from immediate imitating informed short sellers' trading so that we are more able to measure the outcome of informed short sales in our empirical setting.[1]

For our analysis, we build a sample of daily publicly disclosed short sales in Germany before and during the pandemic. In a first step, using a multivariate regression model, we assign healthcare expertise to short sellers that outperform in holding short positions or, alternatively, in covering short positions both in healthcare stocks over the course of seven years reasonably earlier before the outbreak of Cocid-19. In our subsequent main analysis, we date the outbreak event of the Covid-19 pandemic on January 3, 2020 and find that after this date non-healthcare stocks perform worse if their short sellers possess healthcare expertise based on holding short positions or covering them. Since we assume that those outperforming short sellers use more likely public information on Covid-19 for their trades, we interpret this finding as empirical evidence that their information advantage stems from superior information processing skills rather than the use of private information on specific stocks. We find that the value of healthcare expertise is also economically significant: in the case of, for instance, the 10-day cumulative abnormal returns, short sellers with healthcare expertise outperform their control group post-shock by 4.3 percent when using the general shorting performance, and by 7.2 percent when applying the covering performance to identify healthcare expertise. Consistently, we find qualitatively similar outperformances for 5-day and 20-day CAR windows.

We choose German data in our study for several reasons: First, in Germany as in all other EU-member states, short sellers are obliged to disclose changes to their short holdings in listed firms if they exceed 0.5 percent of shares outstanding (e.g., Jan et al. (2019)). Combining these public disclosures with stock data, the high level of detail in our sample allows us to track every position individually

[1] Jank, Roling, and Smajlbegovic (2019) show that short sellers that trade below the public disclosure threshold of 0.5 percent of stocks shorted outperform above-this-threshold short sellers. They argue that disclosure might deter short selling because others might react to them timely potentially causing lower profits.

Covid Economics 65, 20 January 2021: 108-152

throughout the complete sample period from Nov. 1, 2012, through June 30, 2020. Second, Germany has an internationally relevant and sufficiently large publicly-listed healthcare industry. Third, Germany is the only Euro-zone country with developed financial markets that refrains from issuing a ban on short sales during Covid-19. Fourth, even if the UK stock market is larger than the German stock market, we expect that Brexit-induced market volatility and information flow makes it difficult to apply a shock event-based treatment model.

As a first robustness test, we also run our regressions with a sample also including and one only including healthcare stocks. Our results hold for the sample including all industries that shows that our analysis without healthcare stocks does not suffer from a selection bias. In the case of a sample containing solely healthcare stocks, we find no relation of treated short positions to performance that supports our assumption that private information on healthcare stocks such as information on the development of a vaccine or medical treatment against Covid-19 provides no additional information advantage in trading during the Covod-19 pandemic. In addition, we find statistically significant evidence that during the pandemic healthcare expertise is related to higher outperformance in short selling of non-healthcare stocks compared to short selling of healthcare stocks.

In addition, one major concern might be that outperformance of short sellers with healthcare expertise is also driven by factors other than our suggested information-based trading advantage through healthcare expertise during the Covid-19 pandemic. Alternative sources of such outperformance might be the ability to secure funding by short sellers' own investors during pandemic-caused financial market turmoil when fund investors tend to withdraw money, or the ability to locate stocks for borrowing in those volatile pandemic times when stock lenders tend to recall lent-out stocks to trade themselves. To rule out such non information-based trading advantages as alternative explanations for our findings, we conduct several empirical tests.

First, we examine if the outperformance of healthcare expertise short sellers is caused solely by general short selling skills that we expect to stem from the ability to secure funding and locate stocks more likely than the ability to acquire and process information. Applying a measure of general shorting skills, we are not able to replicate our results so that our assumptions are not weakened by alternative explanations such as non information-based skills.

Second, we apply an alternative measure of healthcare expertise that is based on short sellers' fraction of long positions in healthcare stocks retrieved from 13F filings with the SEC. Then, we obtain

Covid Economics 65, 20 January 2021: 108-152

qualitatively the same results. Since this measure is unrelated to technical short selling abilities such as locating stocks, this finding does not support such an alternative explanation.

Third, we retrieve stock lending data from IHS Markit to include lending fee, active utilization, and short selling risk calculated according to Engelberg, Reed, and Ringgenberg (2018) to control for short selling constraints. Again, our findings remain robust to their inclusion so that non informational trading advantages are not supported to drive our results.

Our paper contributes to several strands of the short sale literature. Our findings extend the understanding about the sources of short sellers' information advantage contributing to studies that assign such advantage to the use of private information (e.g., Christophe et al. (2004); Karpoff and Lou (2010)) or public information processing (e.g., Kandel and Pearson (1995); Engelberg et al. (2012)). To our knowledge, we are the first to document a causal relationship for superior information processing skills being a driver for short selling outperformance.

Our additional analysis with 13F data to identify industry-specific trading expertise contributes to the literature on investor skills beyond the narrower view on short sellers (e.g., Kacperczyk and Seru (2007), Cremers and Petajisto (2009), Baker et al. (2010)).

Moreover, our studies contribute to literature on institutional investors during economic crises (e.g., Kacperczyk, Van Nieuwerburgh, and Laura Veldkamp (2011)), in particular during the Covid-19 pandemic (e.g., Pástor and Vorsatz (2020)).

Regarding methodology, we use the Covid-19 outbreak as appropriate identification strategy and thus add to other studies using Covid-19-induced governmental measures for similar identification and difference-in-differences approaches (e.g., Heggeness (2020); Coibion, Gorodnichenko, and Weber (2020); Betcherman et al. (2020); Giommoni and Loumeau (2020)). Most related to our idea that the Covid-19 outbreak is linked to an information advantage for trading and is used for an identification strategy, Henry, Plesko, and Rason (2020) document that insiders of U.S. firms with operations in China trade more frequently early after the Covid-19 outbreak that they attribute to their early access to information on Covid-19 compared to their U.S. counterparts without Chinese operations.

The remainder of this paper is structured as follows: Section 2 provides a literature review and develops the hypothesis. Section 3 describes the data and the empirical strategy. Section 4 presents

Covid Economics 65, 20 January 2021: 108-152

summary statistics and main results. Section 5 includes various robustness tests. Concluding comments are provided in Section 6.

2. Related Literature and Hypothesis

2.1 Related Literature

Short Sellers are generally perceived to be informed and sophisticated traders because a wide range of researchers show that short selling predicts future stock returns (e.g., Asquith et al. (2005); Boehmer et al. (2008); Desai et al. (2002); Diether et al. (2009)). Theoretical models suggest they trade on private information, thereby revealing parts of their information to uninformed investors or copycat traders, and eventually this mechanism causes the information to be incorporated into stock prices (Glosten and Milgrom (1985); Kyle (1985)). Empirical studies find that short selling indeed aids the process of price discovery and improves market efficiency (Aitken et al. (1998), Boehmer and Wu (2013)).

Considerable efforts have been devoted to understanding short sellers' information advantage throughout the last decades. Overvaluation is regularly ascribed as the main motivation for short sellers. Studies find that short sellers have private information about earnings and fundamentals (Boehmer et al. (2020)), and trade on temporary deviation from those fundamentals (Diether et al. (2009)). They are adept at identifying stock-specific overvaluations and avoid shorting undervalued stocks (Boehmer, Huszar, and Jordan (2010)). There are, however, other motivations such as tax, hedging and arbitrage, of which arbitrage seems to be the most prevalent (Brent et al. (1990); Asquith et al. (2005)). The informational content of short sales depends on the underlying motivation, as, for example, arbitrage and hedging trades exert comparably weaker negative impact on stock prices (Aitken et al. (1998)). Moreover, different trader characteristics are associated with different degrees of informational content. Boehmer et al. (2008) show that among individual, institutional and proprietary traders, nonprogram institutional traders' positions are most negatively associated with future stock returns.

Building on the question whether active fund management adds value to investors, a wide body of literature examines if active fund managers possess skills. While the average mutual fund does not outperform passive investment strategies net of fees, many studies find that a small subgroup of mutual funds persistently outperforms (e.g., Kacperczyk and Seru (2007); Cremers and Petajisto (2009); Baker et al. (2010)). Findings are similar for hedge funds in the way that only a subset of traders persistently outperforms (Jagannathan, Malakhov and Novikov (2010); Grinblatt et al.

Covid Economics 65, 20 January 2021: 108-152

(2020)). However, there is strong consensus that short sellers, on average, have an information advantage over other market participants (Asquith et al. (2005); Boehmer et al. (2020); Boehmer et al. (2008); Desai et al. (2002); Diether et al. (2009)).

When asking where the advantage stems from, prevalent explanations are the use of private information or superior processing skills of public information (e.g., Kandel and Pearson (1995)). Agarwal et al. (2013) use quarterly hedge funds' 13F filings to demonstrate outperformance in confidential holdings, suggesting the use of private information. Moreover, managers with a lower reliance on public information perform better than their peers (Kacperczyk and Seru (2007)), and short sellers are shown to trade before the public revelation of financial misrepresentation (Karpoff and Lout (2010)).

On the contrary, Engelberg et al. (2012) examine short sales around news event and find no evidence in favor of private information. They do, however, find evidence for better pulblic information processing skills as short sellers increase trading directly after the publication of news. A skilled information processor converts new public data into valuable trading information, e.g., by analyzing corporate news (Engelberg (2008)). Boehmer et al. (2020) find empirical evidence for short sellers' trading on public superior processed as well as private information. In addition, they document that the information advantage more likely stems from the use of private information.

2.2 Hypothesis

Following literature consensus, informed short sellers obtain superior performance through the use of private information or superior public information processing skills. As outlined above, only few studies address the question as to what extent outperformance stems from either of these sources, as they are difficult to distinguish under normal market conditions.

Contributing to this question, we aim to exploit the unique market conditions during the Covid-19 pandemic to disentangle public information processing skills from private information. The pandemic constitutes a large exogeneous shock to global financial markets that alters market conditions. Since the Covid-19 pandemic is a healthcare crisis by nature, we argue that healthcare-related information which is used to be industry-specific becomes value-relevant for all industries: Firm-specific information (i.e., private information) becomes subordinate to pandemic information that shows global impact and becomes promptly publicly available. Kacperczyk et al. (2011) argue that aggregate payoff shocks are more volatile, and the price of risk is increased during downturns. Fol-

Covid Economics 65, 20 January 2021: 108-152

lowing this, we assume acquiring and processing information about the aggregate impact of a pandemic shock to be more valuable than doing so for micro-level (i.e., firm-level) information: Acquisition of firm-specific information loses its relevance as markets are driven by Covid-19 news which are publicly available in a timely manner due to the global communication infrastructure and public and press interest. Furthermore, we argue that private information value is low in times of the pandemic indicated by the fact that firms' management themselves are not able to assess the impact of Covid-19 as documented by the numerous withdrawals of earnings guidance.[2] So, we expect that short sellers that are experienced in the healthcare industry (henceforth denoted as expertise traders) have an edge over their peers, as the Covid-19 shock enables them to use their industry-specific trading expertise for market-general trading.[3]

In terms of aggregate information processing most relevant for the market level, expertise traders possess knowledge about models of infectious diseases and their applications (e.g., Anderson R.M., Anderson B. and Might (1992); Hethcote (2000); Kermack and McKendrick (1927)), enabling them to forecast global contagion. They can assess probability and severity of lockdowns, shutdowns, stay-at-home orders, and other governmental measures. On the firm level, they are advantaged at assessing winners[4] and losers[5] of Covid-19, or finding resilient and vulnerable geographical regions in regard to their healthcare infrastructure[6], ultimately affecting the workforce of local companies and consumer demand. One might argue that the act of acquiring aggregate healthcare information is the same for expertise and non-expertise short sellers as they belong to the most informed and

Covid Economics 65, 20 January 2021: 108-152

[2] For an overview of withdrawals, see Ashwell, Ben (2020): How Covid-19 is affecting earnings guidance and dividend payments, URL: https://www.irmagazine.com/reporting/how-covid-19-affecting-earnings-guidance-and-dividend-payments, [Oct 31, 2020]

[3] This wording is similar to the terms specific and general human capital in the personnel and labor economics literature.

[4] One example of a Covid-19 winner is HelloFresh AG, a Germany-based company that provides online food services. Driven by increased business during the lockdown period, HelloFresh customer demand more than doubled over the course of the pandemic. Expertise traders might have advantages at assessing the duration and severity of lockdowns.

[5] An example of a Covid-19 loser is TUI AG, a multinational travel and tourism company headquartered in Germany. Lockdowns and travel restriction pose a severe limitation to business activities. Expertise traders might have advantages at assessing the pandemic situation at major destinations and itineraries.

[6] The Spanish healthcare system, for instance, is generally perceived to be of high quality. Nevertheless, Spain became the worst hit European country regarding confirmed infected patients and ranks top three for deaths in Europe (as of Oct. 31, 2020. Reported by John Hopkins University). Expertise traders might have advantages at assessing the pandemic development in Spain which ultimately impacts firms that draw workforce from Spanish regions or engage in business activities with such firms.

sophisticated traders, but nevertheless we expect the pace differs: the processing for non-expertise traders takes longer, but they might achieve the same trading-relevant information in the end.

Which short sellers then possess healthcare expertise? Prior studies define mutual fund or hedge fund manager skills as the ability to persistently generate alpha over a longer time period, e.g., 3 years, 5 years, or even 10 years (e.g., Cremers and Petajisto (2009); Grinblatt et al. (2020), Jagannathan et al. (2010); Baker et al. (2010)). We follow the literature and define healthcare expertise as industry-specific, persistent outperformance in healthcare stocks pre-Covid-19. Thus, we identify the outperforming short sellers pre-Covid-19 and assess their performance during the Covid-19 pandemic. As outlined above, since we assume that private information on healthcare stocks is unlikely material to improve short selling performance in non-healthcare stocks, only information processing skills remain as potential source of information advantage of healthcare expertise in trading non-healthcare stocks. Since we argue that Covid-19 related information is public by nature and can be processed better by healthcare expertise traders, we suggest that short sellers with healthcare expertise profit from superior Covid-19-related public information processing when trading non-healthcare stocks because for those stocks Covid-19 information is also highly relevant. This reasoning leads to our main hypothesis:

> *Healthcare expertise is associated with superior short selling performance in non-healthcare stocks during the Covid-19 pandemic.*

3. Methodology

3.1 Sample Construction

We use German data on publicly disclosed short positions as provided by the German Federal Financial Supervisory Authority[7] (BaFin) from Nov. 1, 2012, through June 30, 2020. Changes in short positions must be disclosed via the Federal Gazette (Bundesanzeiger). We use these disclosures to construct a panel of daily short positions in German stocks, including stocks from direct neighboring countries for which the main trading venue lies within Germany.[8]

[7] In German: Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)
[8] List of stocks and short sellers are tabulated in Appendix A1.

Covid Economics 65, 20 January 2021: 108-152

Starting Nov. 1, 2012, the BaFin implements a two-tier transparency system for disclosing net short positions in stocks exceeding a certain threshold, as constituted by the EU Short Selling Regulation.[9] A first notification to the BaFin must be made by 3:30 p.m. on the following trading day if the net short position exceeds 0.2 percent of a firm's issued shares, and subsequently for each additional 0.1 percent.[10] Upon exceeding 0.5 percent, short sellers are also required to publicly disclose those short positions in the Federal Gazette. Position changes in between two thresholds (e.g., between 0.5 percent and 0.6 percent) are not subject to disclosure. The regulation applies to all issues for which the main trading venue lies within the EU and includes information on position size, stock issuer, ISIN and the name of the investor.[11]

After the outbreak of the Covid-19 pandemic, six EU member states issued a ban on short selling beginning in late April.[12] With Germany and UK not participating in a ban, short selling in UK financial markets might be related to Brexit-induced information. Thus, we argue that Germany is the only country within the EU that provides high-quality daily data on short selling during the pandemic, has developed financial markets, and has a significant healthcare industry with international relevance. Thus, we observe only German data in our study.

Due to the disclosure threshold, we do not observe the exact day of the opening or covering of short positions. Instead, we observe the first day on which the aggregate short position initially surpasses the reporting limit (henceforth opening), the day on which the position declines below the reporting limit (henceforth covering), and changes in between both dates if they exceed another 0.1 percent threshold. As a consequence, a short seller might build up a position before we are able to observe an opening, and keep a position below the threshold after we observe a covering. Consequently, our sample only accounts for positions that are publicly observable, and we are not able to correctly

Covid Economics 65, 20 January 2021: 108-152

[9] See Article 5(1) and 5(2), and Article 6(1) and 6(2) of *Council Regulation (EU) No 236/2012 of 14 March 2012 of the European Parliament and of the Council on short selling and certain aspects of credit default swaps [2012] OJ L86/1.*
[10] On Mar. 16, 2020, the European Securities and Markets Authority (ESMA) adopted a decision to lower the initial reporting threshold to 0.1 percent instead of 0.2 percent. See *European Securities and Markets Authority Decision (EU) No 2020/525 of 16 March 2020 to require natural or legal persons who have net short positions to temporarily lower the notification thresholds of net short positions in relation to the issued shares capital of companies whose shares are admitted to trading on a regulated market above a certain threshold to notify the competent authorities in accordance with point (a) of Article 28(1) of Regulation (EU) No 236/2012 of the European Parliament and of the Council [2020] OJ L116/5.*
[11] Investors can be both natural and legal persons.
[12] Austria, Belgium, France, Greece, Spain and Italy issued a temporary ban on short selling, starting on Mar. 17 or 18, 2020.

Covid Economics 65, 20 January 2021: 108-152

estimate the returns of the entire short position. Jank et al. (2019) use confidential data on short positions below the disclosure threshold to provide insights on the behavior and performance of secretive short sellers.

We build two disjunct data samples: First, we create a larger sample prior to the pandemic (henceforth Training Set[13]) which we use to estimate persistent outperformance in trading healthcare stocks, i.e., healthcare expertise. Second, we create a 1-year sample around the Covid-19 shock (henceforth Covid Set[14]) to apply a difference-in-differences approach, including 6 months of data in the pre-Covid-19 and post-Covid-19 period each. Midnight on June 30, 2019, constitutes the separating date between both sets. The Training Set then covers Nov. 1, 2012, through June 30, 2019, whereas the Covid Set covers Jul. 1, 2019, through June 30, 2020.

In total, our complete sample comprises 266 different short sellers and 214 different stocks in total, including a range of well-known brokers and hedge funds like J.P. Morgan or Renaissance Technologies. Of that, 136 short sellers and 146 stocks show at least one active short position during Covid-19. Throughout the complete sample period, stocks and short sellers appear in 1497 unique combinations,[15] of which the Covid Set covers 38%. Overall, short sellers file 22,313 disclosures in the Training Set and 4,853 in the Covid Set, resulting in 232,757 and 53,010 days with active short positions, respectively.

We use industry classifications from Capital IQ to label each stock as healthcare or non-healthcare. If a business is too diversified or allocates only few resources to healthcare-related activities, we cannot be certain whether traders execute their short sales due to healthcare-related information. Therefore, a stock is classified as healthcare only if most of the firm's business engages in healthcare activities.

We complement our data with stock-level information from the Capital IQ database (e.g., market capitalization, spread, turnover, volatility) and daily abnormal returns. Following prior literature, we implement a conservative approach and assume that sophisticated traders exercise their trades near market close (Jank et al. (2019)). Hence, we estimate returns based on the stock's daily adjusted close price in excess of the dividend-adjusted German Prime All Share Index. This accounts

[13] Referred to as Estimation Window in Event Study Terminology
[14] Referred to as Observation Window in Event Study Terminology
[15] In some cases, a combination of a short seller and stock appears multiple times. This is due to entities that trade an issue through different investment vehicles at the same time. Common examples for distinguishable multi-vehicle entities are Blackrock, Marshall Wace, and Citadel.

for the circumstance that only the aggregate short position at the end of the day (midnight) must be disclosed and intraday changes in position size are generally unobservable. Furthermore, we control for aggregate short interest in a stock using all simultaneously disclosed short positions from the Federal Gazette.

3.2 Empirical Strategy

Training Set: Determining health expertise

Following our hypothesis, we aim to identify a group of short sellers that possess healthcare expertise before the pandemic outbreak. As stated above, expertise refers to the ability to generate persistent abnormal returns in a specific industry (i.e., industry-specific skill). Thus, we estimate the following fixed-effects panel regression model in our Training Set for each short seller *j* individually:

$$Abnormal\ Returns_{i,t} = \alpha_t + \beta_0\left(Position_{i,t}\ x\ Healthcare_j\right) + \gamma_{i,t}X_{i,t} + \delta_i + \theta_t + \varepsilon_{i,t}\ ,$$

where the dependent variable is the 10-day abnormal return in excess of the German Prime All Share Index.[16] $X_{i,t}$ denotes control variables on the firm level *i*.[17] δ_i denotes fixed effects on the firm level. θ_t denotes time fixed effects.

First, we examine the impact of the individual position size on 10-day abnormal returns (henceforth Shorting Expertise). $Position_{i,t}\ x\ Healthcare_i$ denotes the interaction term of individual position size and healthcare stock classification, and the corresponding coefficient is used as the expertise measure. The literature consensus states that a high degree of short selling predicts lower future returns (e.g., Asquith et al. (2005); Boehmer et al. (2008)). Hence, the coefficient is expected to be negatively linked to abnormal returns (i.e., negative coefficient for larger short selling performance) if a short seller is truly skilled. The measure accounts for the overall short selling performance, but we cannot assess if the traders are adept at timing their trades.

We use a variety of stock-specific and trade-specific controls in our models. In particular, we control for position size and opening days; in the spirit of Boehmer et al. (2018), we include coverings days and prior day short interest as reported in the Federal Gazette. Moreover, we include market capitalization (denoted as *MarketCap*), market-to-book ratio, the logarithm of the average prior 5-day

[16] All stock returns here and forthcoming are winsorized at the 1 and 99 level.
[17] On a short seller level, it is difficult to observe specific characteristics, e.g., assets under management or portfolio turnover, since most traders inherent a secretive behavior and data availability on the long side of trades is not as good as on the short side.

Covid Economics 65, 20 January 2021: 108-152

spread plus one, the logarithm of the 3-month volatility plus one, the 1-year beta and the average 5-day turnover. Also, 5-day prior abnormal returns are used to control for momentum effects.[18]

As a second measure to identify healtcare expertise, we implement a measure based on covering of short positions that reflects more precisely timing ability (henceforth Covering Expertise) inspired by Boehmer et al. (2018) as follows:

$$Abnormal\ Returns_{i,t} = \alpha_t + \beta_0\left(Cover_{i,t}\ x\ Healthcare_j\right) + \gamma_{i,t}X_{i,t} + \delta_i + \theta_t + \varepsilon_{i,t}$$

Boehmer et al. (2018) shed light on the price movement around covering dates and find positive stock returns on the exact day of covering, although partially reverting over the course of the next 7 days due to market impact reversal. Their evidence suggests that some short sellers exhibit timing in covering shorts. Likewise, we examine abnormal returns around coverings in healthcare stocks and expect positive returns post-covering if traders are truly skilled. By using 10-day abnormal returns we ensure that a technical price run-up, induced by the market impact of large buy orders, has already reverted by the time we take measure. This way we account for the information-driven impact of coverings. $Cover_{i,t}\ x\ Healthcare_i$ denotes the interaction term of covering days and healthcare stock classification, and β_0 then constitutes our expertise measure.

The samples are reduced to active short positions to reduce noise. Consequently, we apply the models for all 266 short seller and retrieve a list of statistically significant interaction coefficients at the five percent level. Traders are assigned to the healthcare expertise group or control group depending on their measure's association with 10-day abnormal returns and if their corresponding coefficients are statistically significant at the five percent level. Then, we assign short sellers to the healthcare expertise group if the measure for Shorting Expertise (Covering Expertise) is negative (positive).

For those coefficients that are statistically insignificant, or if we lack short disclosures in healthcare stocks, we cannot assume a group assignment due to the disclosure threshold of 0.5 percent in a firm's outstanding shares. Jank et al. (2019) use confidential data provided by the BaFin to show how secretive short sellers consistently refrain from crossing the reporting threshold. So even if we do not observe traders' individual positions in some cases, they might still have a smaller active short position and exert unobservable expertise.[19]

[18] An overview of the variables can be found in Appendix A2.
[19] Example regressions for the expertise measurement are reported in Appendix A3.

Covid Economics 65, 20 January 2021: 108-152

Covid-Set: difference-in-differences approach

Subsequently, we examine the expertise traders' performance post-shock in market-general trading. A naive approach suggests examining the difference in 10-day abnormal returns in non-healthcare stocks for the pre-Covid-19 and post-Covid-19 period. This method, however, captures the entire effect of the Covid-19 pandemic. To isolate the higher value of healthcare expertise, we propose a difference-in-differences approach. Specifically, the expertise group is affected the Covid-19 shock and the shock of elevated expertise value, whereas the control group is exposed to the Covid-19 shock only. If returns are driven by processing of public information, we expect the expertise group to outperform the control group in trading non-healthcare stocks post-Covid-19.

$$Effect\ of\ Healthcare\ Expertise\ = \left(\bar{\alpha}^{post}_{HC\ Value} - \bar{\alpha}^{post}_{HC\ no\ Value}\ \mid\ E = 1\right),$$

To account for the Covid-19 shock, we need to know how the expertise group would have performed given there was no higher value added for healthcare expertise ($\bar{\alpha}^{post}_{HC\ no\ Value}$) . The delta to the observed performance ($\bar{\alpha}^{post}_{HC\ Value}$) then precisely represents the isolated effect of healthcare expertise on performance. However, $\bar{\alpha}^{post}_{HC\ no\ value}$ is unobservable by design. E denotes the expertise group.

Instead, the difference-in-differences approach circumvents unobservable outcomes by imposing parallel trends. That is, had the trading advantage of healthcare expertise not happened, both groups would have shown the same change in average performance. We thus use the observed change in the control group to estimate the unobserved change in the expertise group as follows:

$$Effect\ of\ Healthcare\ Expertise = (\ \bar{\alpha}^{post}_{Exp} - \bar{\alpha}^{pre}_{Exp}\ ) - (\ \bar{\alpha}^{post}_{Control} - \bar{\alpha}^{pre}_{Control}\ ),$$

where $\bar{\alpha}^{post}_{Exp}$ and $\bar{\alpha}^{pre}_{Exp}$ denote the expertise group's average investment performance after and before the shock, respectively. Vice versa, $\bar{\alpha}^{post}_{Control}$ and $\bar{\alpha}^{pre}_{Control}$ denote the control group's average investment performance before and after the shock.

Covid Set: regression model

Identically to our Training Set estimations, 10-day abnormal returns are used as dependent variable in the difference-in-differences approach. We argue that a 10-day return window is sufficiently short

Covid Economics 65, 20 January 2021: 108-152

to capture relevant information in efficient financial markets, and sufficiently long not to be biased by the market impact of large orders.[20]

To apply the difference-in-differences approach, we specify our fixed effects panel regression model as follows:

$$Abnormal\ Returns_{i,j,t} = \alpha_t + \beta_0\left(Covic_t\ x\ Expertise_j\right) + \beta_1 Covid_t + \gamma_{i,j,t} X_{i,j,t} + \delta_{i,j} + \theta_t + \varepsilon_{i,j,t}$$

We implement a dummy variable $Covid_t$ for the post-Covid period and the interaction term $Covic_t\ x\ Expertise_j$ captures the effect of Covid-19 on expertise traders after the shock. The Covid Set ranges from Jul. 1, 2019, through June 30, 2020, and both the pre-Covid-19 and post-Covid-19 period represent a 6-month interval around our suggested shock date on January 3, 2020.

When does the pandemic shock take place?

The choice of shock date is critical. On January 3, 2020, the AAAS[21] is the first association to feature a scientific article about the pandemic on ScienceMag.org[22], thereby providing public and easily accessible information on the virus. Hence, we argue that short sellers possess knowledge about Covid-19 by January 3, 2020. Considering major indices movements, another approach suggests Saturday, Feb. 22, 2020, as shock date. On that day Italy reported the first European Covid-19 casualty,[23] followed by severe declines in global financial markets after the weekend. We refrain from using this shock date as it represents the point in time that the entirety of non-sophisticated traders is unanimously informed about the virus. Alike, we also refrain from using the WHO's declaration of the pandemic outbreak[24] as it is too late to capture relevant effects.[25]

[20] We also test 5-day and 20-day abnormal returns and obtain qualitatively the same results.
[21] American Association for the Advancement of Science
[22] Normile, Dennis (2020): Novel human virus? Pneumonia cases linked to seafood market in China stir concern; URL: https://www.sciencemag.org/news/2020/01/novel-human-virus-pneumonia-cases-linked-seafood-market-china-stir-concern [Oct 31, 2020].
[23] See Mackenzie, James (2020): First Italien patient dies of coronavirus: Ansa news agency; URL: https://www.reuters.com/article/us-china-health-italy-death/first-italian-patient-dies-of-coronavirus-ansa-news-agency-idUKKBN20F2W5 [Oct 31, 2020].
[24] See Farge, Emma, and Michael Shields (2020): World Health Organization calls coronavirus outbreak 'pandemic' for first time; URL: https://www.reuters.com/article/us-health-coronavirus-who-idUSKBN20Y2OI [Oct 31, 2020].
[25] Henry at al. (2020) choose January 19, 2020 as the starting point of the Covid-19 period when the first trading is after the US has begun screening travelers from the Chinese city Wuhan.

Covid Economics 65, 20 January 2021: 108-152

Finally, can the shock occur prior to January 3, 2020? We do not have plausible evidence on public information on the pandemic before our main shock date, however, short sellers might possess indications about Covid-19 at an earlier point in time. In this case, our empirical results would be weakened in statistical significance rendering our result more conservative.

Identifying assumptions

For identification, the parallel trends assumption represents a necessary condition. We assume parallel trends as

$$\mathbb{E}\left[\bar{\alpha}^{post}_{HC\,no\,Value} - \bar{\alpha}^{pre}_{pre\,Shock} \mid E = 1\right] = \mathbb{E}\left[\bar{\alpha}^{post}_{HC\,no\,Value} - \bar{\alpha}^{pre}_{pre\,Shock} \mid E = 0\right],$$

where E = 1 denotes the expertise group and E = 0 denotes the control group. This equation imposes the same difference in average investment performance for both groups from pre-Covid-19 to post-Covid-19 had there been no leverage of healthcare expertise. However, $\bar{\alpha}^{post}_{HC\,no\,Value}$ is an unobserved counterfactual and therewith parallel post-trends are an assumption by design. Instead, we test for parallel pre-trends to ascertain the validity of our difference-in-differences model. Healthcare stocks are excluded from the test as intuition suggests they enable healthcare expertise traders to achieve systematic outperformance. This, however, prevents us from examining the effect of processing skills on market-general trading.

Appendix A4 reports the outcome of our test for parallel pre-trends. Using the 6-month period before the Covid-19 outbreak to identify short-run parallel trends, we find statistically insignificant differences in monthly performance after controlling for group-specific trends. We therefore fail to reject parallel pre-trends (see Kahn and Lang (2019); Roth (2018)). Furthermore, both groups show similar associations with control variables and we assume the groups to be equally affected by other exogeneous forces post-Covid-19 (see Dimick and Ryan (2014), Ryan et al. (2015)).

4. Empirical Results

4.1. Descriptive Statistics

Figure 1 shows the time trend of disclosed shorting activity in the Covid Set. On any given day, an average of 191 positions are public as reported by the Federal Gazette. These positions account for an average aggregate value EUR 52,590 million in short sales.

Noticeably, Graph A reveals interesting changes in shorting activity. Firstly, the total number of short positions that are public stays within a narrow corridor between 230 and 196 active disclosures,

Covid Economics 65, 20 January 2021: 108-152

suggesting there is a rather stable group of traders that consistently cross the reporting threshold (see Jank et al. (2019)). Secondly, starting in mid-October, the total value shorted shows a run-up until the crash in March without noteworthy increases in nominal positions. This spread is related to a simultaneous run-up in stock prices and increasing short interest. At peak, traders hold EUR 66,398 million in public short sales. Lastly, when the downward price movements commence in mid-March, active short positions increase and decrease inverse to market returns, but only as the markets already plunge, suggesting short sellers capitalize on the momentum.

Graph B underlines March as the busiest month for short sellers. We observe an average of 41 openings and coverings each month except March, which shows twice the activity.

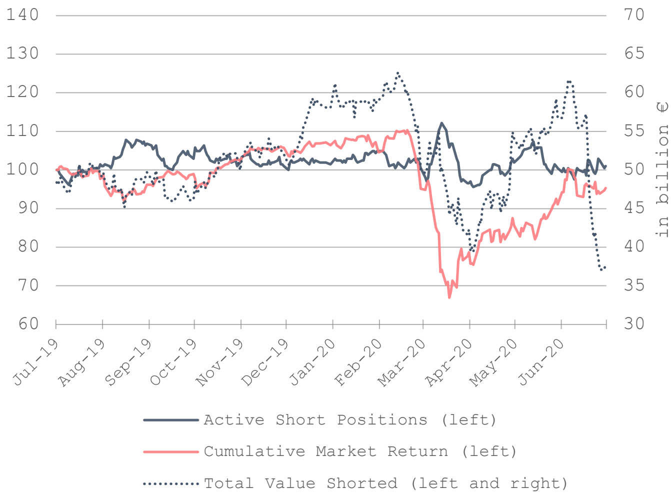
Covid Economics 65, 20 January 2021: 108-152

Figure 1

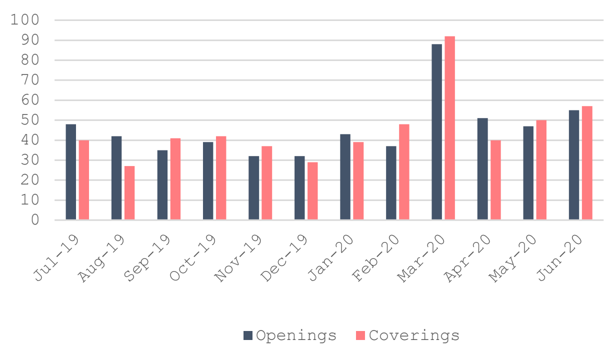
Timeline of Active Short Positions, Total Value Shorted and Market Returns

Graph A in Figure 1 shows the time series of publicly disclosed short positions compared to the market benchmark. All time series variables are relative to their respective level on Jul. 1, 2019 (left axis). *ActiveShortPositions* denote the nominal amount of positions that simultaneously exceed the reporting threshold. *TotalValueShorted* denotes the aggregate value of all public short positions (absolute values on right axis). *CumulativeMarketReturn* is based on the German Prime All Share Index.

Graph A. Relative Amount of Shorts Positions, their Aggregate Value, and Market Returns



Graph B. Monthly Openings and Coverings



Covid Economics 65, 20 January 2021: 108-152

Both the day with the most openings and the day with the most coverings take place during March.[26]

Detailed summary statistics are reported in Table 1. We report statistics both with (A) and without (B) healthcare stocks, but we do find not a significant difference in magnitudes for any of the variables. To classify the characteristics of the Covid Set, we refer to related studies by Jank et al. (2019) who use the same disclosure mechanism in a sample with German stocks. Levels in short interest and marker capitalization are very similar, however, spread is only one half to one fourth of the levels observed by Jank et al. (2019), suggesting enhanced liquidity during Covid-19 markets. Overall, statistical characteristics are in similar ranges. When comparing to Boehmer et al. (2018) who use disclosure rules on Japanese markets to assess covering trades, characteristics differ to a greater extent, i.e., market capitalization, market to book, and aggregate short interest is lower for Japanese stocks. The difference in short interest is likely due to the sample construction by Boehmer et al. (2018) including all listed stocks, whereas we only examine stocks that have at least one active short position. The differences in the remaining variables pertain to the lower threshold of 0.25 percent for public disclosure in Japan, making more stocks eligible for the list of stocks with publicly held short positions, and there are more small firms than large firms. The disparity suggests that lowering the threshold forces more public short disclosures for mid or small caps, ultimately reducing the average and median market capitalization of shorted stock.

As expected, we observe a negative mean and median for abnormal returns, except for the 20-day median. Excluding healthcare stocks yields even lower returns which coincides with the intuition that healthcare stocks might not be as affected by the crisis as other firms might be. When disaggregating summary statistics (reported in Table 2) into pre-Covid-19 and post-Covid-19 intervals, we find that mean and median returns are invariably positive prior to Covid-19 and negative afterwards. Interestingly, the level of aggregate short interest is rather stable which coincides with the observed narrow corridor of nominal public disclosures.[27] We observe slightly greater averages for market capitalization which might be due to more short selling in large firms, volatility and spread increase post-Covid-19.

Covid Economics 65, 20 January 2021: 108-152

[26] Maximum openings (16) on March 9. Maximum coverings (9) on March 24.
[27] See Figure 1. Active disclosures are between 196 and 230 at any given point.

CEPR PRESS

Covid Economics 65, 20 January 2021: 108-152

Table 1

Descriptive Statistics of Publicly Disclosed Short Positions from BaFin amid the Covid-19 pandemic from Jul. 2019 to June 2020 (Covid Set)

Table 1 reports detailed summary statistics on the Covid Set with (A) and without healthcare stocks (B), active short positions only. *Ln ShortInterest$_{-1}$* is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover$_{[-5,-1]}$* is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread$_{-1}$* is the prior day difference of bid and ask divided by their average. *Momentum$_{[-5,-1]}$* is the cumulative abnormal return of the five preceding trading days. *Ln Indicative Fee* is the annualized indicative lending fee for a stock. *Rebate Rate* is the annualized rebate rate. *Short Risk* is the natural logarithm of 1-year lending fee volatility. *Ln Active Utilization* is the percentage of shares out on loan relative to available shares for lending. Returns are in excess of the German Prime All Share daily returns.

A: Covid Set

Variable	N	Mean	Median	Std. Dev.	Min	Max
BaFin						
Ln Short Interest$_{-1}$	52,675	1.09	1.24	0.97	-0.76	2.80
Short Interest$_{-1}$	52,805	4.36	3.46	3.41	0.00	16.39
Position Size	53,010	1.03	0.75	0.78	0.01	9.80
Capital IQ						
Ln MarketCap	53,010	7.47	7.48	1.18	0.67	11.7
MarketCap	53,010	3,573	1,768	5,986	1.96	120,000
Market-To-Book	52,566	1.60	1.18	1.09	0.63	11.78
Ln Spread$_{[-5,-1]}$	52,802	-1.51	-1.65	0.89	-4.15	1.74
Spread$_{[-5,-1]}$	52,802	0.34	0.19	0.38	0.02	5.68
Turnover$_{[-5,-1]}$	51,752	0.62	0.39	1.40	0.00	36.43
Ln Volatility	52,822	0.93	0.93	0.42	-0.50	3.02
Volatilty	52,822	2.76	2.54	1.15	0.60	20.40
Beta 1-Year	52,934	1.34	1.29	0.72	-1.24	3.57
Momentum$_{[-5,-1]}$	52,967	-0.06	-0.08	5.80	-19.02	17.51
Markit IHS						
Ln Indicative Fee	40,881	0.55	0.24	1.28	-1.01	3.15
Indicative Fee	40,881	3.98	1.27	5.42	0.36	23.38
Rebate Rate	40,881	-2.79	-0.48	5.62	-22.66	1.92
Short Risk	52,492	3.55	3.65	1.65	0.01	7.49
Ln Active Utilization	40,895	3.32	3.68	1.26	-1.07	4.61
Active Utilization	40,895	47.27	39.79	36.17	0.34	100.00
Abn. Returns 5-days	53,010	0.02	-0.02	5.95	-19.02	17.51
Abn. Returns 10-days	53,010	0.03	-0.07	8.53	-27.15	23.64
Abn. Returns 20-days	53,010	-0.01	0.10	12.04	-36.00	34.41

B: Covid Set without Healthcare Stocks

Variable	N	Mean	Median	Std. Dev.	Min	Max

BaFin						
Position Size	48,309	1.03	0.76	0.01	0.75	9.80
Ln Short Interest$_{-1}$	47,998	1.07	0.95	-0.76	1.21	2.80
Short Interest$_{-1}$	48,120	4.25	3.36	0.00	3.33	16.39
Capital IQ						
Ln MarketCap	48,309	7.43	1.19	0.67	7.38	11.7
MarketCap	48,309	3,451	5,633	2	1,605.49	120,000
Market-To-Book	47,979	1.49	0.96	0.63	1.16	11.78
Ln Spread$_{[-5,-1]}$	48,117	-1.48	0.89	-4.15	-1.60	1.74
Spread$_{[-5,-1]}$	48,117	0.35	0.38	0.02	0.20	5.68
Turnover$_{[-5,-1]}$	47,182	0.63	1.46	0.00	0.38	36.43
Ln Volatility	48,121	0.94	0.43	-0.50	0.94	3.02
Volatility	48,121	2.80	1.18	0.60	2.56	20.40
Beta 1-Year	48,309	1.35	0.74	-1.24	1.31	3.57
Momentum$_{[-5,-1]}$	48,269	-0.07	5.82	-19.02	-0.10	17.51
Markit IHS						
Ln Indicative Fee	37,235	0.58	1.29	-1.01	0.29	3.15
Indicative Fee	37,235	4.13	5.56	0.36	1.34	23.38
Rebate Rate	37,235	2.94	5.76	-1.92	0.48	22.66
Short Risk	47,806	3.64	1.62	0.01	3.69	7.49
Ln Active Utilization	37,241	3.34	1.26	-1.07	3.73	4.61
Active Utilization	37,241	47.80	35.87	0.34	41.48	100.00
Abn. Returns 5-days	48,309	-0.01	6.03	-19.02	-0.05	17.51
Abn. Returns 10-days	48,309	-0.05	8.65	-27.15	-0.10	23.64
Abn. Returns 20-days	48,309	-0.16	12.20	-36.00	0.07	34.41

Covid Economics 65, 20 January 2021: 108-152

Table 2

Descriptive Statistics of Publicly Disclosed Short Positions Pre- and Post-Covid (Covid Set)

Table 2 reports detailed summary statistics for the Covid Set (active Positions only, without healthcare stocks) for the pre-Covid and post-Covid sample window. *Ln ShortInterest*$_{-1}$ is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover*$_{[-5,-1]}$ is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread*$_{-1}$ is the prior day difference of bid and ask divided by their average. *Momentum*$_{[-5,-1]}$ is the cumulative abnormal return of the five preceding trading days. *Ln Indicative Fee* is the annualized indicative lending fee for a stock. *Rebate Rate* is the annualized rebate rate. *Short Risk* is the natural logarithm of 1-year lending fee volatility. *Ln Active Utilization* is the percentage of shares out on loan relative to available shares for lending. Returns are in excess of the German Prime All Share daily returns.

	Pre Covid		Post Covid	
Variable	**Mean**	**Median**	**Mean**	**Median**
BaFin				
Ln Short Interest$_{-1}$	1.11	1.24	1.07	1.28
Short Interest$_{-1}$	4.33	3.44	4.4	3.59
Position Size	1.04	0.77	1.02	0.73
Capital IQ				
Ln MarketCap	7.45	7.45	7.5	7.51
MarketCap	3,461	1,717	3,688	1,819.02
Market-To-Book	1.55	1.19	1.64	0.96
Ln Spread$_{[-5,-1]}$	-1.64	-1.81	-1.38	-1.45
Spread$_{[-5,-1]}$	0.3	0.16	0.38	0.24
Turnover$_{[-5,-1]}$	0.45	0.32	0.79	0.47
Ln Vola 6 months	0.78	0.83	1.09	1.11
Vola 6 months	2.32	2.28	3.21	2.35
Beta 1-Year	1.48	1.49	1.19	0.77
Momentum$_{[-5,-1]}$	-0.05	-0.09	-0.07	-0.07
Markit IHS				
Ln Indicative Fee	0.42	0.02	0.62	-0.62
Indicative Fee	3.42	1.02	4.3	1.56
Rebate Rate	-1.35	1	-3.64	-0.96
Short Risk	3.46	3.57	3.65	3.76
Ln Active Utilization	3.15	3.46	3.43	3.9
Active Utilization	41.67	31.72	50.59	49.57
Abn. Returns 5-days	0.11	0.06	-0.06	-0.12
Abn. Returns 10-days	0.14	-0.01	-0.09	-0.22
Abn. Returns 20-days	0.26	0.31	-0.29	-0.08

Covid Economics 65, 20 January 2021: 108-152

CEPR PRESS

Covid Economics 65, 20 January 2021: 108-152

4.2 Main Results

Table 3 displays the results of our difference-in-differences estimation. The interaction terms for Shorting Expertise represent statistically significant alphas of -2.0 percent, -4.3 percent, and -7.7 percent for 5-day, 10-day, and 20-day stock returns, respectively. They reflect the average alpha of expertise traders over their non-expertise control group during the pandemic indicating that the expertise group indeed exhibits superior performance. Moreover, from 5-day to 10-day returns and from 10-day to 20-day returns, the alpha grows 109 percent and 81 percent, respectively. This coefficient growth indicates that outperformance is persistent within our measured return windows, and expertise traders increase their alpha as their holding periods lengthen. It seems the effect is more pronounced for the 5-day to 10-day transition and shows deceleration for longer return-windows.

Specifications (4)-(6) show the regression results for Covering Expertise. Consistent with the results for Shorting Expertise, we find significant alphas of -2.6 percent, -5.9 percent, and -10.8 percent for 5-day, 10-day and 20-day returns, respectively.

Consistent with our hypothesis, our findings indicate that expertise traders possess superior processing skills in Covid-19 markets, as healthcare information is then also relevant for non-healthcare stocks. Taken together, since short sellers devote their resources to processing Covid-19 information, the expertise traders have an edge through their superior processing skills of healthcare-related information, and ultimately generate alpha over their control group with no healthcare expertise in market-general trading.

The coefficients on the *Covid* variable in specifications (1)-(6) reflect the expected mean change in returns from pre- to post-shock for the control group. There are only statistically significant coefficients for the Covering Expertise that are positive indicating a worse short selling performance by non-expertise short sellers post-Covid-19.

Table 3
Main Results

In Table 3, we report the Difference-in-Differences estimation for healthcare expertise traders and non-expertise traders using the Covid Set (active positions, without healthcare stocks). Group affiliation is determined via Shorting Expertise or Covering Expertise from the Training Set. *Covid X Expertise* and *Covid* denote the variables of interest for the Difference-in-Differences estimation. *Covid* is a dummy variable that equals 1 if the date is later or equal to Jan. 3, 2020. *Covid X Expertise* is the interaction term of *Covid* and the *Expertise* dummy. *Ln ShortInterest*$_{-1}$ is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover*$_{[-5,-1]}$ is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread*$_{-1}$ is the prior day difference of bid and ask divided by their average. *Momentum*$_{[-5,-1]}$ is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively.

	Shorting Expertise			**Covering Expertise**		
	Abnormal Returns			Abnormal Returns		
Variables	5 days	10 days	20 days	5 days	10 days	20 days
	(1)	(2)	(3)	(4)	(5)	(6)
Covid x Expertise	**-2.040****	**-4.268****	**-7.746*****	**-3.355*****	**-7.163*****	**-13.159*****
	(0.814)	(1.628)	(2.718)	(0.900)	(1.852)	(3.355)
Covid	0.091	1.080	2.517	2.196**	5.686***	12.205***
	(0.935)	(1.063)	(1.631)	(1.000)	(1.591)	(2.915)
Ln Short Interest$_{-1}$	2.004***	3.800***	6.522***	1.074**	2.024**	3.397*
	(0.592)	(1.135)	(1.978)	(0.455)	(0.906)	(1.748)
Position Size	-0.085	-0.036	-0.185	0.165	0.308	0.773
	(0.469)	(0.893)	(1.466)	(0.318)	(0.652)	(1.170)
Opening	-0.382	-0.675	0.952	0.047	0.299	2.483**
	(0.801)	(0.995)	(1.374)	(0.740)	(0.934)	(1.130)
Covering	0.452	1.075	1.610*	0.836*	1.135*	1.180
	(0.686)	(0.672)	(0.878)	(0.464)	(0.581)	(0.790)
Ln MarketCap	-4.265***	-10.083***	-20.069***	-4.545***	-10.926***	-22.563***
	(0.847)	(1.627)	(2.686)	(0.908)	(1.622)	(2.685)
Market-To-Book	0.922*	1.839*	5.113***	1.020**	2.002*	5.898**
	(0.514)	(1.001)	(1.850)	(0.507)	(1.167)	(2.379)
Ln Spread$_{[-5,-1]}$	-0.278	-1.571**	-3.827***	-0.262	-1.323*	-2.679**
	(0.472)	(0.781)	(1.204)	(0.432)	(0.754)	(1.128)
Turnover$_{[-5,-1]}$	-0.025	-0.327	-0.864**	-0.142	-0.446*	-1.287***
	(0.194)	(0.231)	(0.389)	(0.182)	(0.238)	(0.358)
Ln Vola 6m	0.713	1.057	1.126	1.538	2.925	3.844
	(1.381)	(2.701)	(4.370)	(1.193)	(2.320)	(3.734)
Beta 1 year	-0.408	-1.078	-2.189	-0.091	-0.690	-1.621
	(0.390)	(0.807)	(1.431)	(0.380)	(0.743)	(1.305)
Momentum$_{[-5,-1]}$	0.007	0.016	0.035*	0.005	0.001	0.012

Covid Economics 65, 20 January 2021: 108-152

CEPR PRESS

	(0.010)	(0.015)	(0.018)	(0.007)	(0.010)	(0.014)
Constant	29.364***	69.688***	136.931***	31.079***	75.846***	156.835***
	(6.202)	(12.132)	(19.942)	(6.779)	(12.100)	(19.973)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Short Seller - Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,560	8,560	8,560	12,939	12,939	12,939
R-squared	0.037	0.074	0.146	0.038	0.076	0.148

5. Robustness Tests

5.1 Alternative Explanation: private Information on healthcare stocks

One concern with our analysis might be that our results suffer from a selection bias because we only look at non-healthcare stocks. Even though we do this on purpose to rule out private information-based Covid-19 treatment effects, we also run our main regression with a larger sample including healthcare stocks and obtain the same qualitive results as shown in Table 4.

If we run our regressions on a sample only containing healthcare stocks, we do not estimate any relation of treated short positions to abnormal returns. Tests for differences in coefficients also reveal that our effect is more pronounced for non-healthcare stocks. If the use of private information were also an alternative source of short sale outperformance, we would expect to see a treatment effect of Covid-19 on healthcare stocks as well. Since we do not observe this, we regard this as support for our public information processing assumption behind the treatment effect.

Covid Economics 65, 20 January 2021: 108-152

Covid Economics 65, 20 January 2021: 108-152

Table 4

Robustness: Healthcare Stocks

In Table 4, we vary the main analysis by including or excluding healthcare stocks in the Covid Set. Group affiliation is determined via Shorting Expertise or Covering Expertise from the Training Set. *Covid X Expertise* and *Covid* denote the variables of interest for the Difference-in-Differences estimation. *Covid* is a dummy variable that equals 1 if the date is later or equal to Jan. 3, 2020. *Covid X Expertise* is the interaction term of *Covid* and the *Expertise* dummy. *Ln ShortInterest*$_{-1}$ is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover*$_{[-5,-1]}$ is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread*$_{-1}$ is the prior day difference of bid and ask divided by their average. *Momentum*$_{[-5,-1]}$ is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

A: Shorting Expertise

Variables	No Healthcare	Only Healthcare	Full Sample	Difference	Difference	Difference
	(1)	(2)	(3)	(1) - (3)	(2) - (3)	(1) - (2)
Covid x Expertise	**-4.268****	**0.380**	**-4.153*****	**-0.115**	**4.533***	**-4.648****
	(1.628)	(1.599)	(1.506)	**(0.850)**	**(0.0523)**	**(0.0354)**
Covid	1.080	1.520	1.012			
	(1.063)	(3.314)	(1.080)			
Constant	69.688***	133.275***	67.335***			
	(12.132)	(20.777)	(11.889)			
Controls	Yes	Yes	Yes			
Time FE	Yes	Yes	Yes			
Short Seller - Stock FE	Yes	Yes	Yes			
Observations	8,560	1,015	9,575			
R-squared	0.074	0.278	0.069			

B. Covering Expertise

Variables	No Healthcare	Only Healthcare	Full Sample	Difference	Difference	Difference
	(1)	(2)	(3)	(1) - (3)	(2) - (3)	(1) - (2)
Covid x Expertise	**-7.163*****	**-4.551**	**-7.046*****	**-0.117**	**2.495**	**-2.612**
	(1.852)	(2.287)	(1.719)	**(0.6934)**	**(0.3396)**	**(0.3454)**
Covid	5.686***	3.120	5.408***			
	(1.591)	(4.513)	(1.541)			

Constant	75.846*** (12.100)	172.833*** (39.229)	77.713*** (11.704)
Controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Short Seller - Stock FE	Yes	Yes	Yes
Observations	12,939	719	13,658
R-squared	0.076	0.313	0.070

5.2 Alternative explanation: non information-based trading advantage

One major concern might be that outperformance of short sellers with healthcare expertise is also driven by factors other than our suggested information-based trading advantage. Alternative sources of such outperformance might be the ability to secure funding by short sellers' own investors during pandemic-caused financial market turmoil when fund investors tend to withdraw money, or the ability to locate stocks for borrowing in those volatile pandemic times when stock lenders tend to recall lent-out stocks to trade themselves. To rule out such non information-based trading advantages as alternative explanations for our findings, we conduct several empirical tests.

General shorting skills

First, we examine if the outperformance of healthcare expertise short sellers is driven only by general short selling skills that we expect to be highly correlated with the ability to secure funding and locate stocks. So, we re-estimate *Expertise* in the regression for the Training Set without the Interaction with the *Healthcare*. Consistent with our previous applied methodology, we assign short sellers to the expertise group according to the coefficients on *Position* and *Covering* for Shorting Expertise and Covering Expertise, respectively. Applying those new measured expertise assignments to our regressions of the Covid Set (exhibited in Table 5), we obtain no statistically significant findings. So, we conclude that general shorting skills do not drive our results so that our information-based trading advantage is not weakened.

Covid Economics 65, 20 January 2021: 108-152

Table 5

Robustness: General Shorting Skills

In Table 5, we report outcomes of robustness tests using General Shorting Skills, i.e., Shorting Expertise and Covering Expertise from short selling of all stocks instead of only healthcare stocks, in the Covid Set (active positions, without healthcare stocks). *Covid X Expertise* and *Covid* denote the variables of interest for the Difference-in-Differences estimation. *Covid* is a dummy variable that equals 1 if the date is later or equal to Jan. 3, 2020. *Covid X Expertise* is the interaction term of *Covid* and the *Expertise* dummy. *Ln ShortInterest$_{-1}$* is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover$_{[-5,-1]}$* is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread$_{-1}$* is the prior day difference of bid and ask divided by their average. *Momentum$_{[-5,-1]}$* is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Variables	**Shorting Expertise**	**Covering Expertise**
	(1)	(2)
Covid x Expertise	**3.339***	**-0.238**
	(1.690)	(1.432)
Covid	-0.850	-0.079
	(1.324)	(0.954)
Ln Short Interest$_{-1}$	0.802	-1.024
	(0.879)	(0.693)
Position Size	-0.739	3.582***
	(1.951)	(0.757)
Opening	-2.790	-0.899
	(1.924)	(1.144)
Covering	0.616	1.622**
	(1.610)	(0.671)
Ln MarketCap	-11.473***	-15.344***
	(2.695)	(2.313)
Market-To-Book	-0.794	1.735
	(1.908)	(2.043)
Ln Spread$_{[-5,-1]}$	-0.595	-0.408
	(0.844)	(0.828)
Turnover$_{[-5,-1]}$	-0.636**	-0.447
	(0.265)	(0.402)
Ln Vola 6m	-2.804	-3.993
	(2.769)	(2.567)
Beta 1 year	-0.037	-0.374
	(1.374)	(1.143)
Momentum$_{[-5,-1]}$	0.005	0.009
	(0.016)	(0.015)
Constant	85.768***	114.072***

Covid Economics 65, 20 January 2021: 108-152

	(17.911)	(15.907)
Time FE	Yes	Yes
Short Seller - Stock FE	Yes	Yes
Observations	7,650	9,421
R-squared	0.069	0.093

Healthcare expertise based on long positions

To further support that our treatment group is influenced solely by information-related to Covid-19 through processing skills, we identify healthcare expertise if short sellers' long positions in healthcare stocks according to their 13F filings with the SEC is above the median of the sample average over two years preceding the Covid Set. Since this approach is not based on short selling performance and we assume that long investing is not necessarily correlated short selling performance, we assume that this approach rules out the impact of non information-based trading advantages regarding short selling.

As exhibited in Table 6, we find qualitatively the same results as found with our short selling data-based measures of healthcare expertise so that our results hold even if we rule out a important fraction of non information-based trading advantages.

Covid Economics 65, 20 January 2021: 108-152

Table 6

Robustness: Alternative 13F Healthcare Expertise

In Table 6, we report outcomes of robustness tests using 13F filings as alternative measure for healthcare expertise. We compute the percentage of 13F portfolios allocated to healthcare stocks using data from IHS Markit. Short sellers are denoted as expertise traders in specification (1) if their healthcare allocation exceeds the median of all 13F allocations. *Covid X Expertise* and *Covid* denote the variables of interest for the Difference-in-Differences estimation. *Covid* is a dummy variable that equals 1 if the date is later or equal to Jan. 3, 2020. *Covid X Expertise* is the interaction term of *Covid* and the *Expertise* dummy. *Ln ShortInterest*$_{-1}$ is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover*$_{[-5,-1]}$ is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread*$_{-1}$ is the prior day difference of bid and ask divided by their average. *Momentum*$_{[-5,-1]}$ is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Variables	Full Sample	< $Q_{0.25}$	$Q_{0.25}$-$Q_{0.5}$	$Q_{0.5}$-$Q_{0.75}$	> $Q_{0.75}$
	(1)	(2)	(3)	(4)	(5)
Covid x Expertise	**-2.024***	**0.002**	**3.260***	**-1.761***	**-0.680**
	(0.889)	(0.821)	(1.343)	(0.758)	(0.756)
Covid	2.454***	1.029**	0.506	1.570***	1.178**
	(0.737)	(0.460)	(0.507)	(0.480)	(0.487)
Ln Short Interest$_{-1}$	1.339***	1.450***	1.282***	1.330***	1.464***
	(0.468)	(0.463)	(0.467)	(0.465)	(0.464)
Position Size	0.540	0.613	0.481	0.629	0.587
	(0.495)	(0.521)	(0.484)	(0.518)	(0.523)
Opening	-0.385	-0.351	-0.410	-0.345	-0.357
	(0.538)	(0.540)	(0.540)	(0.538)	(0.540)
Covering	1.056**	1.105**	1.037**	1.072**	1.104**
	(0.434)	(0.438)	(0.438)	(0.436)	(0.439)
Ln MarketCap	-8.649***	-8.371***	-8.533***	-8.532***	-8.335***
	(0.896)	(0.892)	(0.905)	(0.889)	(0.887)
Market-To-Book	0.308	0.224	0.114	0.332	0.182
	(0.845)	(0.833)	(0.824)	(0.856)	(0.829)
Ln Spread$_{[-5,-1]}$	-0.462	-0.486	-0.480	-0.449	-0.480
	(0.378)	(0.382)	(0.375)	(0.375)	(0.379)
Turnover$_{[-5,-1]}$	-0.350***	-0.341***	-0.355***	-0.345***	-0.339***
	(0.112)	(0.113)	(0.112)	(0.112)	(0.113)
Ln Volatility	0.891	1.059	1.037	0.974	1.042
	(1.119)	(1.143)	(1.121)	(1.120)	(1.152)
Beta 1 year	-1.030***	-0.990***	-1.074***	-0.928**	-1.002***
	(0.379)	(0.375)	(0.380)	(0.377)	(0.376)
Momentum$_{[-5,-1]}$	0.003	0.003	0.003	0.003	0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)

Covid Economics 65, 20 January 2021: 108-152

Constant	62.371***	60.029***	61.846***	61.173***	59.878***
	(6.475)	(6.578)	(6.498)	(6.407)	(6.516)
Time FE	Yes	Yes	Yes	Yes	Yes
Short Seller - Stock FE	Yes	Yes	Yes	Yes	Yes
Observations	46,519	46,519	46,519	46,519	46,519
R-squared	0.050	0.048	0.052	0.050	0.049

We do not take this measure as our main measure although we would get more observations because long positions are less costly than short positions so that internal incentives to invest in expertise for those long positions is lower than for short positions that might question the value of the expertise measured. In addition, long positions might be matched to unobserved short selling activity that would also question our information-based treatment effect assumption. Moreover, quarterly reporting of long positions is too infrequent to measure expertise that is used for much shorter investment horizons in the case of short selling.

Stock lending data

For further support our assumption that information drives the treatment effect, we also control for short selling conditions and constraints as potential alternative driver of non information-based trading advantages. So, we include lending fee and active utilization that we retrieve on a daily basis from IHS Markit. In addition, we include short selling risk calculated according to Engelberg et al. (2018). As shown in Table 7, our results remain qualitatively the same that indicates that our results are not driven by an important fraction of non information-based trading advantages.

Covid Economics 65, 20 January 2021: 108-152

Table 7

Stock Lending Data

In specification (1) and (2) of Table 7, we report probit regressions on variables denoting short selling constraints. *Covid* is a dummy variable that equals 1 if the date is after or equal to Jan. 3, 2020. *Ln Indicative Fee* is the annualized indicative lending fee for a stock. *Short Risk* is the natural logarithm of 1-year lending fee volatility. *Ln Active Utilization* is the percentage of shares out on loan relative to available shares for lending. *Ln ShortInterest*$_{-1}$ is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover*$_{[-5,-1]}$ is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread*$_{-1}$ is the prior day difference of bid and ask divided by their average. *Momentum*$_{[-5,-1]}$ is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Variables	Shorting Expertise	Covering Expertise
	(1)	(2)
Covid x Expertise	-3.816*	-5.898**
	(2.100)	(2.261)
Covid	0.844	4.636***
	(1.150)	(1.768)
Ln Indicative Fee	0.738	-1.454*
	(0.861)	(0.834)
Short Risk	-0.877	0.190
	(0.796)	(0.565)
Ln Active Utilization	-0.345	0.438
	(0.342)	(0.490)
Ln Short Interest$_{-1}$	4.103***	2.303**
	(1.268)	(1.042)
Position Size	0.055	-0.146
	(0.850)	(0.693)
Opening	-0.597	0.415
	(1.159)	(1.073)
Covering	1.121	0.682
	(0.773)	(0.692)
Ln MarketCap	-10.775***	-12.029***
	(1.890)	(1.898)
Market-To-Book	1.431*	2.007
	(0.858)	(1.363)
Ln Spread$_{[-5,-1]}$	-1.831**	-1.144
	(0.837)	(0.903)
Turnover$_{[-5,-1]}$	-0.350	-0.486*
	(0.308)	(0.272)
Ln Volatility	1.137	2.516
	(2.882)	(2.411)
Beta 1-Year	-1.939**	-1.372
	(0.745)	(0.852)
Momentum$_{[-5,-1]}$	0.031*	0.008
	(0.017)	(0.013)

Covid Economics 65, 20 January 2021: 108-152

Constant	79.341***	84.954***
	(15.067)	(13.958)
Time FE	Yes	Yes
Firm FE		
Short Seller - Firm FE	Yes	Yes
Observations	6,463	9,617
R-squared	0.081	0.074

6. Conclusion

We use public German data on daily short sales before and during the Covid-19 pandemic to isolate the processing skills of healthcare expertise traders from the use of private information and assess if better information processing through healthcare expertise is associated with superior returns.

Consistent with our hypothesis, we find overwhelming evidence that healthcare expertise causes market-wide outperformance in the Covid-19 pandemic. We argue that short sellers shift their resources to gather and process aggregate pandemic-induced information and private, firm-specific information becomes subordinate in the fast-moving and volatile Covid-19 markets. The formerly industry-specific healthcare expertise then enables expertise traders to outperform. Hence, our findings are in line with the studies by Engelberg et al. (2018) and provide causal evidence that processing skills are a main driver of returns. Overall, we offer new insights into the success of short sales. We show that superior processing skills is a driver of performance, and to our knowledge, we are the first to document a causal relationship for this.

Our research is relevant to market participants and researchers to understand the competitive edge of some short sellers and the origin of their success. Form the regulators' point of view, our results support that short sellers might be good information processors and therefore are important for improving market efficiency by using public (legal) information (e.g., Saffi and Sigurdsson (2011)). We also document how industry-specific expertise becomes general expertise under certain conditions that provides a new perspective on labor economics' notion of specific human capital.

Traders show varying motivations for short selling, such as hedging, arbitrage, tax reasons or overvaluation. It would be interesting for future research to differentiate expertise traders across their investment strategies for further insights. Moreover, our data is limited to the Federal Gazette publications, but as the BaFin confidentially collects short positions below that threshold, we suggest

Covid Economics 65, 20 January 2021: 108-152

that future researchers with access investigate if secretive expertise traders exert even better information processing, and if the disclosure rules constitute a binding short constraint that impairs expertise trader performance. It might also be fruitful to examine if the outperformance of expertise traders repeats in the second wave of Covid-19 where short selling constraints rise again. If the outperformance does not repeat, it is more likely evidence for our suggested information-based trading advantage that other short sellers then also experience when have invested in healthcare expertise on their own.

References

Agarwal, V., Jiang, W., Tang, Y. and Yang, B. "Uncovering Hedge Fund Skill from the Portfolio Holdings they hide." *Journal of Finance*, 68 (2013), 739–783.

Aitken, Michael J., Alex Frino, Michael S. McCorry and Peter L. Swan. "Short Sales Are Almost Instantaneously Bad News: Evidence from the Australian Stock Exchange." *The Journal of Finance*, 43 (1998), 2205-2223.

Anderson, Roy M., and Robert M. May. "Infectious Diseases of Humans: Dynamics and Control." *Oxford University Press*, 1991, 12-23.

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter. "Short Interest, Institutional Ownership, and Stock Returns." *Journal of Financial Economics*, 78 (2005), 243-276.

Baker, Malcolm, Lubomir Litov, Jessica Wachter, and Jeffrey Wurgler. "Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements." *Journal of Financial and Quantitative Analysis*, 45 (2010), 1111–1131.

Betcherman, Gordon, N. Giannakopoulous, I. Laliotis, I. Pantelaiou, M. Testaverde, and G. Tzimas. "Reacting quickly and protecting jobs: The short-term impacts of the COVID-19 lockdown on the Greek labor market." Covid Economics, 43 (2020), 95-136.

Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu. "Short-Sale Constraints, Differences of Opinion, and Overvaluation." *Journal of Financial and Quantitative Analysis*, 41 (2006), 455-487.

Boehmer, Ekkehart, Truong X. Duong, and Zsuzsa R. Huszár. „Short Covering Trades." *Journal of Financial and Quantitative Analysis*, 53 (2018), 723-748.

Boehmer, Ekkehart, Zsuzsa R. Huszar, and Bradford D. Jordan. „The good news in short interest." *Journal of Financial Economics*, 96 (2010), 80-97.

Boehmer, Ekkehart, Charles M. Jones, Juan (Julie) Wu, and Xiaoyan Zhang. „What Do Short Sellers know?" *Review of Finance*, (2020), 1-33.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang. "Which Shorts Are Informed?" *Journal of Finance*, 63 (2008), 491–527.

Brent, Averil, Dalte morse, and E. Kay Stice. "Short Interest: Explanations and Tests.*" The Journal of Financial and Quantitative Analysis*, 25 (1990), 273-289.

Bris, Arturio, William N. Goetzmann, and Ning Zhu. "Efficiency and the Bear: Short Sales and Markets Around the World." *The Journal of Finance*, 62 (2007), 1029-1079.

Christophe, Stephen E., Michael G. Ferri, and James J. Angel. "Short selling prior to earnings announcements." *Journal of Finance*, 59 (2004), 1845–1875.

Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. "*American Journal Of Sociology*, 100 (1995), 1261-1293.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. "Labor Markets during the Covid-19 Crisis: A preliminary View." *National Bureau of Economic Research*, 2020.

Cremers, K. J. Martin, and Antti Petajisto. "How Active is Your Fund Manager? A New Measure That Predicts Performance." *The Review of Financial Studies*, 22 (2009), 3329-3365.

D'aviolo, Gene. "The Market for Borrowing Stock." *Journal of Financial Economics*, 66 (2002), 271-306.

Desai, Hemang, K. Ramesh, S. R. Thiagarajan, and Bala V. Balachandran. "Investigation of the Information Role of Short Interest in the NASDAQ Market." *Journal of Finance*, 57 (2002), 2263–2287.

Diamond, Douglas W., and Robert E. Verrecchia. "Constraints on Short-Selling and Asset Price Adjustment to Private Information." *Journal of Financial Economics*, 18 (1987), 277-311.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner. "It's SHO Time! Short Sale Price Tests and Market Quality." *Journal of Finance*, 64 (2009), 37–73.

Dimick, Justin B., and Andrew M. Ryan. "Methods for Evaluationg Changes in Health Care Policy. The Difference-in-Differences Approach." *Jama*, 312.22 (2014), 2401-2402.

Duffie, Darrell, Nicolae Gârleanu, and Lasse H. Pedersen. „Securities lending, shorting, and pricing." *Journal of Financial Economics*, 66 (2002), 307-339.

Covid Economics 65, 20 January 2021: 108-152

Duong, Truong X., Zsuzsa R. Huszár, and Takeshi Yamada. "The costs and benefits of short sale disclosure." *Journal of Banking & Finance*, 53 (2015), 124-139.

Engelberg, Joseph. "Costly Information Processing: Evidence from Earnings Announcements" *Unpublished Working Paper*, 2008. University of North Carolina.

Engelberg, Joseph, Adam V. Reed, and Matthew C. Ringgenberg. "Short-Selling Risk" *The Journal of Finance*, 73 (2018), 755-786.

Engelberg, Joseph E., Adam V. Reedn, and Metthew C. Ringgenberg. "How are Shorts Informed? Short Sellers, News, and Information Processing". *Journal of Financial Economics*, 105 (2012), 260-278.

Evans, Richard, Miguel A. Ferreira, and Melissa P. Prado. "Fund Performance and Equity Lending: Why Lend What You Can Sell?" *Review of Finance*, 21 (2017), 1093-1121.

Giommoni, Tommaso, and Gabriel Loumeau. "Lockdown and voting behaviour: A natural experiment on postponed elections during the COVID-19 pandemic." *Covid Economics*, 41 (2020), 69 - 99.

Glosten, Lawrence R., and Paul R. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, 14 (1985), 71-100.

Grinblatt, Mark, Gergana Jostova, Lubomir Petrasek, and Alexander Philipov. "Style and Skill: Hedge Funds, Mutual Funds, and Momentum." *Management Science*, (2020), 1-27.

Heggeness, Misty L. "Estimating the immediate impact of the COVID-19 parental attachment to the labor market and the double bind of mothers." *Review of Economics of the Household*, 18 (2020), 1053-1078.

Henry, Erin, George A. Plesko, and Caleb Rawson. "Geographic Proximity and Insider Trading: Evidence from COVID-19." *Available at SSRN 3678121*, 2020.

Hethcote, Herbert W. "The mathematics of Infectious Diseases". *SIAM Review*, 42 (2000), 599-653.

Jank, Stephan, Christoph Roling, and Esad Smajlbegovic, "Flying Under the Radar: The Effects of Short-Sale Disclosure Rules on Investor Behavior and Stock Prices", Journal of Financial Economcis (forthcoming), *Available at SSRN 2806886*, 2019.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov. "Do Hot Hands Exist among Hedge Fund Managers? An Empirical Evaluation." *The Journal of Finance*, 65 (2010), 217-255.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp. “Rational Attention Allocation Over the Business Cycle.” *NBER Working Paper Series*, (2011).

Kacperczyp, Marcin, and Amit Seru. “Fund Manager Use of Public Information: New Evidence on Managerial Skills.” *The Journal of Finance*, 62 (2007), 485-528.

Kahn-Lang, Arielle, and Kevin Lang. „The The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications.” *Journal of Business & Economic Statistics*, 38 (2019), 1537-2707.

Kandel, Eugene, and Neil D. Pearson. “Differential interpretation of public signals and trade in speculative markets.” *Journal of Political Economy*, 103 (1995), 831–872.

Karpoff, Jonathan M., and Xiaoxia Lou. “Short Sellers and Financial Misconduct” *The Journal of Finance*, 65 (2010), 1879-1913.

Kermack, William Ogilvy, and Anderson G. McKendrick. “A contribution to the mathematical theory of epidemics.” *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character*, 115.772 (1927), 700-721.

Kyle, Albert S. “Continuous Auctions and Insider Trading.” *Econometrica*, 53 (1985), 1315-1335.

Miller, Edward M. “Risk, Uncertainty, and Divergence of Opinion.” *The Journal of Finance*, 32 (1977), 1151-1168.

Pástor, Ľuboš, and M. Blair Vorsatz. “Mutual Fund Performance and Flows during the COVID-19 Crisis.” *The Review of Asset Pricing Studies*, 10 (2020), 791-833.

Prado, Melissa P., Padro A. C. Saffi, and Jason Sturgess. “Ownership Structure, Limits to Arbitrage, and Stock Returns: Evidence from Equity Lending Markets.” *The Review of Financial Studies*, 29 (2016), 3211-3244.

Roth, Jonathan. “Should We Adjust for the Test for Pre-trends in Difference-in-Difference Designs?” *arXiv preprint*, 1804.01208 (2018).

Ryan, Andrew M., James F. Burgess Jr, and Justin B. Dimick. “Why we should not be indifferent to specification choices for difference-in-differences.” *Health services research*, 50.4 (2015), 1211-1235.

Saffi, Pedro A. C., and Kari Sigurdsson. “Price Efficiency and Short Selling”, Review of Financial Studies 24(3) (2011), 821–852.

Appendix A1

Table A1

Overview of stocks and short sellers

Table A1 provides a comprehensive list of (1) stocks, (2) short sellers and (3) an overview over the most active short sellers in the Covid Set. ACTIVE DAYS denotes the nominal number of days on which a short position is publicly disclosed (>0.5%). AVG POSITION SIZE is the average over the days on which a position is public, relative to shares outstanding.

(1) Stocks

1&1 Drillisch AG	ENCAVIS AG	PATRIZIA AG
Aareal Bank AG	Evonik Industries AG	ProSiebenSat.1 Media SE
adidas AG	Evotec AG	publity AG
ADLER Real Estate AG	Fielmann AG	PUMA SE
ADVA Optical Networking SE	flatex AG	QIAGEN
AIXTRON SE	Fraport AG	QSC AG
Allianz SE	freenet AG	Rheinmetall AG
alstria office REIT-AG	Fresenius SE & Co. KGaA	RIB Software SE
AROUNDTOWN	GEA Group AG	RTL GROUP SA
Aumann AG	Gerresheimer AG	RWE AG
AURELIUS SE	Gerry Weber International AG	S&T AG
Aurubis AG	GFT Technologies AG	SAF HOLLAND
BASF SE	GRENKE AG	Salzgitter AG
BAUER AG	HAMBORNER REIT AG	SAP SE
Bayer AG	Hamburger Hafen und Logistik	Sartorius AG
Bayerische Motoren Werke AG	HeidelbergCement AG	Schaeffler AG
Bechtle AG	Heidelberger Druckmaschinen	Scout24 AG
Befesa	HELLA GmbH & Co. KGaA	SGL Carbon SE
Beiersdorf AG	HelloFresh SE	SHOP APOTHEKE EUROPE
Bertrandt AG	Henkel AG & Co. KGaA	Siemens AG
Bilfinger SE	HOCHTIEF AG	Siltronic AG
Borussia Dortmund	HUGO BOSS AG	Sixt SE
Brenntag AG	Infineon Technologies AG	SLM Solutions Group AG
CANCOM SE	Instone Real Estate Group AG	SMA Solar Technology AG
Carl Zeiss Meditec AG	ISRA VISION AG	SNP Schneider SE
CECONOMY AG	JOST Werke AG	Software AG
COMMERZBANK AG	Jungheinrich AG	Stabilus AG
Continental AG	K+S AG	Steinhoff International
Core State Capital Holding SA	KION GROUP AG	Ströer SE
Covestro AG	Klöckner & Co SE	Südzucker AG
CTS Eventim AG & Co. KGaA	Knorr-Bremse AG	TAG Immobilien AG
CYAN AG	Koenig & Bauer AG	TeamViewer AG
Daimler AG	KRONES AG	technotrans SE
Delivery Hero SE	LANXESS AG	Tele Columbus AG

Covid Economics 65, 20 January 2021: 108-152

DEUTSCHE BANK AG	LEG Immobilien AG	Telefónica Deutschland Holding AG
Deutsche Börse AG	Leifheit AG	thyssenkrupp AG
Deutsche EuroShop AG	LEONI AG	TOM TAILOR Holding AG
Deutsche Lufthansa AG	LPKF Laser & Electronics AG	TUI AG
Deutsche Pfandbriefbank AG	Medigene AG	Uniper SE
Deutsche Post AG	MERCK KgaA	United Internet AG
Deutsche Telekom AG	METRO AG	va-Q-tec AG
Deutsche Wohnen SE	MorphoSys AG	VARTA AG
DEUTZ AG	MTU Aero Engines AG	Voltabox AG
Dialog Semiconductor	Munich Re	Wacker Chemie AG
DIC Asset AG	Nemetschek SE	Wacker Neuson SE
Drägerwerk AG & Co. KGaA	New Work SE	WashTec AG
Dürr AG	Nordex SE	Wirecard AG
E.ON SE	NORMA Group SE	Zalando SE
ElringKlinger AG	OSRAM Licht AG	zooplus AG

(2) Short Sellers

Adage Capital Mgmt	FourWorld Capital Mgmt	Park West Asset Mgmt
Adehi Capital	GF Trading	PDT Partners
AHL Partners	Gladstone Capital Mgmt	Pelham Capital
AKO Capital	GLG Partners	Petrus Advisers
Albar Capital	GMT Capital Corp	Pictet Asset Mgmt
Amia Capital	Greenvale Capital	Point72 Asset Mgmt
Anchorage Capital Master Offshore	GSA Capital Partners	Polar Capital
AQR Capital Mgmt	Half Sky Capital (UK)	Polygon Global Partners
Arrowstreet Capital	Harbor Spring Capital	Portsea Asset Mgmt
Atom Investors	HBK Investments	PSquared Asset Mgmt
Balyasny Asset Mgmt	Helikon Investments	Public Equity Partners Mgmt
BlackRock	Henderson Global Investors	Qube Research & Technologies
Bloom Tree Partners	Highbridge Capital Mgmt	Renaissance Technologies
BlueCrest Capital Mgmt	Immersion Capital	Rye Bay Capital
BlueMountain Capital Mgmt	Jericho Capital Asset Mgmt	Samlyn Capital
BNP PARIBAS	JPMorgan Asset Mgmt (UK)	Sand Grove Capital Mgmt
BODENHOLM CAPITAL	Kairos	Sandbar Asset Mgmt
Bridgewater Associates	Kintbury Capital	Sanditon Asset Mgmt
Bybrook Capital	Kontiki Capital Mgmt (HK)	Scopia Capital Mgmt
Caledonia (Private) Investments	Kuvari Partners	Sculptor Capital Mgmt Europe
Canada Pension Plan Inv Board	Lakewood Capital Mgmt	Slate Path Capital
CapeView Capital	Lancaster Investment Mgmt	Soros Fund Mgmt
Capital Fund Mgmt	Lansdowne Partners (UK)	Squarepoint Ops
Caxton Associates	Lazard Asset Mgmt	Susquehanna International
Citadel	Leucadia Investment Mgmt	Sylebra Capital
Coatue Mgmt	LMR Partners	Systematica Investments

Covid Economics 65, 20 January 2021: 108-152

Coltrane Asset Mgmt	Lone Pine Capital	TCI Fund Mgmt
Connor Clark & Lunn Inv Mgmt	Makuria Inv Mgmt (UK)	Thames River Capital
Covalis Capital	Maple Rock Capital Partners	Think Investments
CPMG	Maplelane Capital	Third Point
CQS (UK)	Marshall Wace	Thunderbird Partners
Credit Suisse International	Maverick Capital	Tiger Global Mgmt
D. E. Shaw & Co.	MEAG MUNICH ERGO	Tower House Partners
Darsana Capital Partners	Melqart Asset Mgmt (UK)	TT International
Davidson Kempner	Melvin Capital Mgmt	Two Creeks Capital Mgmt
DNB Asset Mgmt	Merian Global Investors (UK)	Tybourne Equity
Duquesne Family Office	Meritage Group	UBS Asset Mgmt
Eleva Capital SAS	Millennium	Valiant Capital Mgmt
Elliott Investment Mgmt	Muddy Waters Capital	Viking Global Investors
Eminence Capital	Naya Capital Mgmt UK	Voleon Capital Mgmt
ENA Investment Capital	No Street	Wellington Mgmt Company
Engadine Partners	Numeric Investors	Whale Rock Capital Mgmt
Ennismore Fund Mgmt	Oceanwood Capital Mgmt	Whitebox Advisors
EXANE ASSET MGMT	Odey Asset Mgmt	Winton Capital Mgmt
ExodusPoint Capital Mgmt	Otus Capital Mgmt	WorldQuant
Fosse Capital Partners	Paloma Partners Mgmt	Zimmer Partners

(3) Short Sellers	Active Days	Avg Position Size
Marshall Wace	4182	0.96
BlackRock	3876	1.23
AQR Capital Mgmt	2944	1.25
Citadel	2810	1.24
Canada Pension Plan Investment Board	1917	0.83
Millennium	1571	0.64
Ennismore Fund Management	1290	1.25
GLG Partners	1213	0.91
JPMorgan Asset Management	1114	0.85

Covid Economics 65, 20 January 2021: 108-152

Appendix A2

Table A2
Variables and Basic Fixed effects Panel Regression Model

In Table A2, we report the basic structure of our fixed-effects panel regression model, used in both the Training Set and the Corona Set. The same structure applies for each analysis. We adjust the Variable of Interest for different interaction terms, depending on the purpose. An example for the application of this model to estimate a short seller's expertise can be found in Appendix C.

Variables	
Abnormal Returns	5-day, 10-day or 20-day abnormal returns in excess of the German Prime All Share Index, based on dividend-adjusted close prices, winsorized at the 1 and 99 level
Covid x Healthcare Expertise	Interaction term and Variable of Interest
Covid Dummy	Denotes the post-shock period, starting from Jan. 3
Healthcare Expertise Dummy	Denotes whether a short seller has healthcare expertise
Healthcare Dummy	Denotes whether a stock is classified as healthcare
Position Size	Individual position size as % of shares outstanding
Opening Dummy	Denotes the exact day of short position opening
Covering Dummy	Denotes the exact day of short position covering
Short Interest$_{-1}$	Prior day aggregate net short positions as % of shares outstanding
Ln MarketCap	Natural Logarithm of same day market capitalization
Market-To-Book	(MarketCap +Total Assets – Book Value of Common Equity) over Total Assets
Ln Spread$_{[-5,-1]}$	Natural Logarithm of 5 to 1-day prior average bid-ask spread
Turnover$_{[-5,-1]}$	5 to 1-day prior average turnover relative to WASO
Ln Volatility	Natural Logarithm of 6-month average volatility
Beta	1-year beta
Momentum$_{[-5,-1]}$	5 to 1-day prior cumulative abnormal returns
Indicative Fee	Indicative Lending Fee p.a.
Rebate Rate	Rebate Rates p.a.
Short Risk	Natural logarithm of 1-year volatility of Indicative Fee
Active Utilization	Percentage of Shares on Loan relative to Available Shares for lending

Covid Economics 65, 20 January 2021: 108-152

Covid Economics 65, 20 January 2021: 108-152

Appendix A3

Table A3

Exemplary Expertise Measurement in Training Set

Table A3 displays an example of expertise measurement for the short seller Millennium Management in the Training Set. The Shorting Expertise variable *Position x Healthcare* is an interaction term of the position size and healthcare stock classification. The Covering Expertise variable *Covering x Healthcare* is an interaction term of a dummy for covering days and healthcare stock classification. *Ln ShortInterest$_{-1}$* is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover$_{[-5,-1]}$* is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread$_{-1}$* is the prior day difference of bid and ask divided by their average. *Momentum$_{[-5,-1]}$* is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively.

	Shorting Expertise	**Covering Expertise**
Variables	Abnormal Returns 10 days	Abnormal Returns 10 days
Position x Healthcare	**-4.829*** **	
	-1.516	
Covering x Healthcare		**7.105*** **
		-2.480
Ln Short Interest$_{-1}$	0.149	0.136
	(0.206)	(0.210)
Position Size	-1.002	-2.822**
	(1.237)	(1.244)
Opening	-0.149	-0.280
	(0.424)	(0.405)
Covering	-0.473	-0.980
	(0.655)	(0.671)
Ln MarketCap	8.103*	8.951*
	(4.292)	(4.739)
Market-To-Book	0.447	0.356
	(1.040)	(1.030)
Ln Spread$_{[-5,-1]}$	2.346**	2.239*
	(1.117)	(1.121)
Turnover$_{[-5,-1]}$	-0.010	-0.037
	(0.769)	(0.773)
Ln Volatility	-20.512***	-21.592***
	(6.414)	(7.011)
Beta 1-Year	-0.538	-0.650
	(0.744)	(0.710)
Momentum$_{[-5,-1]}$	-0.164***	-0.160***
	(0.027)	(0.027)

CEPR PRESS

Constant	-7.711	-9.339
	(17.350)	(18.467)
Time FE	Yes	Yes
Short Seller - Stock FE	Yes	Yes
Observations	3,980	3,980
R-squared	0.2233	0.2198
Short Seller	**Millennium Management**	**Millennium Management**
Classification	**Expertise Group**	**Expertise Group**

Covid Economics 65, 20 January 2021: 108-152

CEPR PRESS

Covid Economics 65, 20 January 2021: 108-152

Appendix A4

Table A4

Test for Parallel Pre-Trends (Covid Set)

In Table A4, we test for monthly parallel pre-trends in the Covid Set. Pre-Trends are indicated by the monthly interaction with the *Expertise* variable. Expertise Group Affiliation is determined via Shorting Expertise or Covering Expertise from the Training Set. *Ln ShortInterest$_{-1}$* is the aggregate of all public short positions relative to shares outstanding. *Position Size* is the size of individual short positions relative to shares outstanding. *Opening* and *Covering* are dummy variables for days with openings and coverings. *Ln MarketCap* is the natural logarithm of market capitalization in EUR millions. *Market-To-Book* is defined in Table 2. *Ln Volatility* is the 6-month volatility. *Beta 1-Year* is the 1-year rolling beta to its closest benchmark. *Turnover$_{[-5,-1]}$* is the average turnover of the five preceding trading days, relative to shares outstanding. *Ln Spread$_{-1}$* is the prior day difference of bid and ask divided by their average. *Momentum$_{[-5,-1]}$* is the cumulative abnormal return of the five preceding trading days. Stock-level data is based on Capital IQ. Returns are in excess of the German Prime All Share daily returns. We include time-fixed and firm-short-seller fixed effects. All Standard errors are clustered robust at the firm-short-seller level and reported in parentheses. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively.

	Shorting Expertise	**Covering Expertise**
Variables	Abnormal Returns 10 days	Abnormal Returns 10 days
July x Expertise	2.182	-0.230
	(2.643)	(1.492)
August x Expertise	2.905	2.715
	(2.707)	(2.350)
September x Expertise	1.510	1.795
	(2.433)	(2.043)
October x Expertise	6.038**	6.919***
	(2.419)	(1.791)
November x Expertise	4.362*	2.211
	(2.469)	(1.604)
Ln Short Interest$_{-1}$	3.964***	1.937**
	(1.163)	(0.935)
Position Size	-0.018	0.615
	(0.921)	(0.708)
Opening	-0.690	0.368
	(1.006)	(0.916)
Covering	1.078	1.184**
	(0.664)	(0.590)
Ln MarketCap	-10.065***	-12.042***
	(1.585)	(1.648)
Market-To-Book	1.625	1.884**
	(1.019)	(0.952)
Ln Spread$_{[-5,-1]}$	-1.553*	-1.256*
	(0.816)	(0.736)
Turnover$_{[-5,-1]}$	-0.347	-0.495**

	(0.234)	(0.243)
Ln Volatility	0.975	2.924
	(2.689)	(2.289)
Beta 1-Year	-1.115	-0.717
	(0.824)	(0.788)
Momentum$_{[-5,-1]}$	0.015	0.002
	(0.015)	(0.010)
Constant	85.768***	114.072***
	(17.911)	(15.907)
Post-Shock Interactions	Yes	Yes
Time FE	Yes	Yes
Short Seller - Stock FE	Yes	Yes
Observations	8,560	12,939
R-squared	0.082	0.096

Covid Economics 65, 20 January 2021: 108-152

Shifts in the portfolio holdings of euro area investors in the midst of COVID-19: Looking through investment funds[1]

Daniel Carvalho[2] and Martin Schmitz[3]

Date submitted: 6 January 2021; Date accepted: 18 January 2021

Covid Economics 65, 20 January 2021: 153-187

We study the impact of the COVID-19 shock on the portfolio exposures of euro area investors. The analysis "looks through" holdings of investment fund shares to rst gauge euro area investors' full exposures to global debt securities and listed shares by sector at end-2019 and to subsequently analyse the portfolio shifts in the rst and second quarter of 2020. We show important heterogeneous patterns across asset classes and sectors, but also across euro area less and more vulnerable countries. In particular, we find a broad-based rebalancing towards domestic sovereign debt at the expense of extra-euro area sovereigns, consistent with heightened home bias. These patterns were strongly driven by indirect holdings – via investment funds – especially for insurance companies and pension funds, but levelled off in the second quarter. On the contrary, for listed shares we find that euro area investors rebalanced away from domestic towards extra-euro area securities in both the first and the second quarter, which may be associated with better relative foreign stock market performance. Many of these shifts were only due to indirect holdings, corroborating the importance of investment funds in assessing investors' exposures via securities, in particular in times of large shocks. We also confirm this mechanism in an analysis focusing on the large-scale portfolio rebalancing observed between 2015 and 2017 during the ECB's Asset Purchase Programme.

1 We thank Philip Lane, Fausto Pastoris and an anonymous reviewer for helpful comments and suggestions. We also thank participants at the BdF-ECB Workshop on international financial flows, representatives of the Eurosystem's Monitoring Working Group and seminar participants at the University of Mainz. The views expressed are those of the authors and do not necessarily reflect those of Banco de Portugal or the European Central Bank.

2 Senior Economist, Banco de Portugal.
3 Lead Economist-Statistician, European Central Bank.

CEPR PRESS

1 Introduction

The unprecedented health measures introduced in the first quarter of 2020 to contain the COVID-19 pandemic unleashed a severe financial shock. The latter had a striking impact on euro area financial markets, with far-reaching, heterogeneous consequences across countries and sectors. In particular, intra-euro area fragmentation dynamics set off in March, on account of safe haven and flight-to-quality behaviour, which affected in particular the foreign holdings of debt securities issued by more vulnerable euro area countries – Greece, Italy, Portugal and Spain – while less vulnerable countries – i.e., Austria, Belgium, Finland, France, Germany and the Netherlands – recorded foreign inflows.

Against this background, this paper studies the portfolio shifts and rebalancing in euro area investors' holdings and exposures in the first two quarters of 2020. In order to achieve this goal, the paper proposes a method of "looking-through" the holdings of investment fund shares by euro area investors, so that a clearer picture of their full exposures is attained.

There are two main reasons why looking through investment funds is a relevant feature. First, from a structural point of view, looking through the holdings of investment funds to assess the exposures of euro area sectors is relevant as some sectors have large holdings of investment fund shares.

Second, euro area investment funds played an important role in the developments in financial flows amidst the COVID-19 shock, with a stark retrenchment in their cross-border holdings in the first quarter of 2020, followed by an almost equally sized rebound in the second quarter.

In broad terms, our analysis shows that investors tend to resort proportionally more to investment fund shares when they invest in listed shares than to debt securities. Moreover, they highlight a high heterogeneity across sectors regarding their direct and indirect exposures structure. On the one hand, a significant share of pension funds', insurance companies' and households' overall exposures are attained via investment fund shares, which for the latter may be the result of the lower level of investor sophistication typically associated to this sector; on the other hand, banks' exposures are mostly via direct holdings, especially when it comes to debt securities. These idiosyncratic differences point to the need to explore the patterns of exposures at the sectoral level instead of at the aggregate level. The latter is further reinforced by the differentiated behaviour that sectors have in response to returns – see, for instance, Adrian and Shin (2010), Adrian and Shin (2013), Timmer (2018) and Bergant and Schmitz (2020) – as well as the need to comply with certain regulatory requirements, which apply to some sectors. [1]

[1] For other studies, which explore sectoral patterns, see, for instance, Bergant et al. (2020), Giofré (2013), Roque and Cortez (2014), Galstyan et al. (2016), Boermans and Vermeulen (2019) and Galstyan and Velic (2018).

Covid Economics 65, 20 January 2021: 153-187

To assess the shifts in exposures of the different sectors in the first and second quarter of 2020, we use a regression approach, which combines a set of gravity variables, standard in the literature, with a battery of geography and issuer sector dummies. There are a number of interesting general points to be made from the results. First, a "reversion to the mean" effect is present, especially in the case of debt securities, whereby the larger exposures at the end of 2019 were the most reduced during the first quarter of 2020. Second, many of the observed portfolio shifts were only due to the component of indirect exposures, corroborating the importance of investment funds in providing exposure of sectors to securities and in adjusting those exposures.

Third, focusing on debt securities and starting with the shifts observed in the first quarter of 2020, there was a widespread re-orientation towards domestic sovereigns at the expense of extra-euro area sovereigns, strongly driven by indirect holdings for some sectors, in both less and more vulnerable euro area countries, in an environment of a general sell-off of the latter securities by foreign investors. The picture is not so-clear cut when it comes to securities issued by banks and non-financial corporations (NFC), with some sectors reducing their exposures to securities issued by foreign banks and increasing to securities issued by NFC, across the different geographies. In the second quarter of 2020, however, these shifts came to a halt and even, in the case of some sectors, investment in non-domestic debt securities resumed, against the backdrop of the stabilisation seen in financial markets.

Fourth, moving to listed shares, euro area investors rebalanced away from domestic towards extra-euro area securities in the first quarter of 2020, which is the opposite of what was observed for debt securities, and was strongly driven by indirect holdings. In contrast to debt securities, these investment patterns largely continued in the second quarter of 2020. What can explain the difference in rebalancing patterns between debt securities and listed shares? On the one hand, the rebalancing from non-euro area securities towards domestic debt securities is consistent with heightened home bias and a reduction in foreign currency risk in times of crises. It is well documented that home bias tends to be higher in debt securities than equities – in part due to foreign currency risk (Fidora et al., 2007). Furthermore, Broner et al. (2014) argue that investors may have an incentive to buy domestic debt to the extent they are positively discriminated vis-à-vis foreign investors, since they have a lower probability of being negatively affected in a default episode; domestic investors may also buy domestic debt due to moral suasion.[2] On the other hand, the rebalancing towards non-euro area listed shares may have been driven by the better relative stock market performance in non-European indices (in particular, in the US) and the fact that, in the second half of the first quarter of 2020, Europe was at the centre of the pandemic's developments

[2] See also Broner et al. (2010) and Altavilla et al. (2017).

Covid Economics 65, 20 January 2021: 153-187

CEPR PRESS

with strong containment measures being enacted. What is more, this behaviour is in line with Broner et al. (2006), who find that mutual funds, when facing returns below average, tend to retrench from stocks of countries in which they are positioned overweight.

In a final exercise, we apply our "look-through" framework to the rebalancing observed during the most intense phase of the ECB's Asset Purchase Programme (APP) from early-2015 to end-2017. The results of this exercise show that euro area investors increased their exposure to securities issued by non-residents, which was heavily driven by indirect holdings for most sectors. Thus, the main vehicle to increase exposures to extra-euro area debt during the APP period, was precisely the one used to reduce these exposures during the COVID-19 shock: indirect holdings via investment fund shares.

The remainder of the paper is organised as follows: Section 2 presents descriptive evidence on the rebalancing patterns by euro area investors in the first and second quarter of 2020. Section 3 presents the methodology used to look through the holdings of investment fund shares and Section 4 shows the descriptive results of the look-through approach. Our empirical approach is introduced in Section 5, the results of the regression exercise are discussed in Section 6; Section 7 focuses on the period of the ECB's APP; finally, Section 8 concludes.

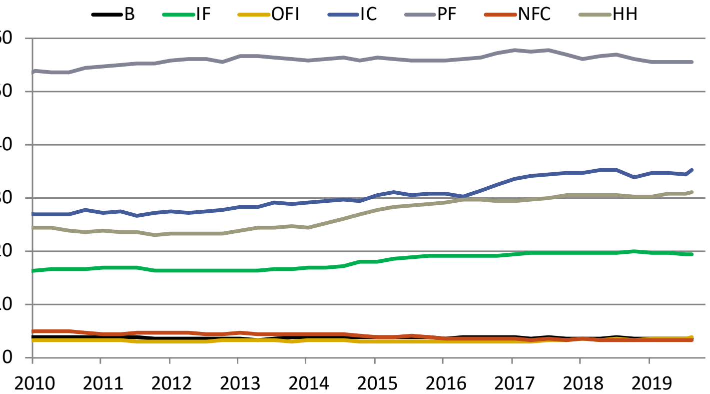
Covid Economics 65, 20 January 2021: 153-187

2 Descriptive evidence on euro area rebalancing amid the COVID-19 pandemic

Euro area investors, in particular some sectors, have large holdings of investment fund shares. Figure 1 displays the proportion of investment fund shares in the aggregate security portfolios of euro area sectors: pension funds are the sector with the highest allocation in investment fund shares, followed by insurance companies and households, as well as investment funds themselves. In contrast, banks, other financial intermediaries and non-financial corporations exhibit relatively low allocations.

Euro area investment funds played an important role in the developments in euro area financial flows amidst the COVID-19 shock. Figures 2a and 2b display Securities Holding Statistics (SHS) data, which, unlike balance of payments data, also include domestic transactions. Starting with the asset side, these show that most of the pro-cyclical deleveraging and the subsequent rebound were carried by investment funds followed, by a long distance, by insurance companies. In terms of instruments, investment funds executed net sales of debt securities, listed shares and investment fund shares in the first quarter of 2020, and made net purchases of these instruments with very similar magnitudes in the second quarter. Banks, on the other hand, had a counterbalancing effect in the first quarter, taking on both foreign and domestic

Figure 1: Proportion of investment fund shares in the aggregate portfolio of euro area sectors



Source: ECB.

Covid Economics 65, 20 January 2021: 153-187

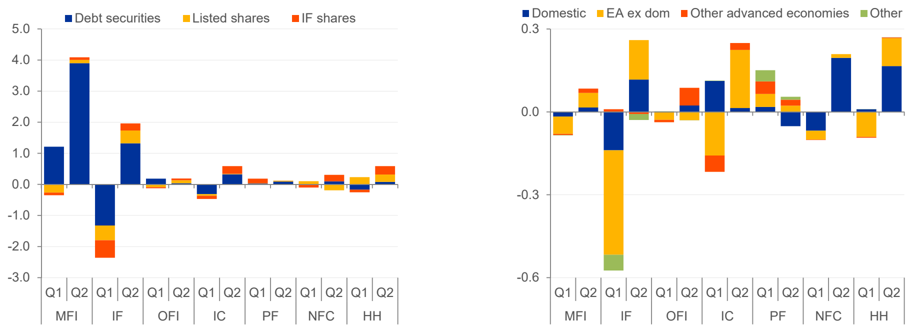
debt securities. In the second quarter these net purchases by banks increased even further.

Moving to the liability side of investment funds, Figure 2b focuses solely on transactions of euro area residents in investment fund shares. It shows that all euro area sectors – with the exception of pension funds – reduced their exposures to other securities via investment funds in the first quarter of 2020, while increasing it again in the second quarter. In turn, this points to the usefulness of trying to determine the indirect exposures via investment funds of other euro area sectors, in order to be able to have a better idea of how these portfolio shifts took place.[3] Figure 2b also shows that the overwhelming portion of the transactions were in investment funds shares issued in the euro area, either domestically or in other euro area countries, mainly in Ireland and Luxembourg. Transactions in non-euro area investment fund shares by euro area investors were mostly residual, which is to be expected, given the low relative exposure of euro area investors to investment fund located outside the euro area.

Monthly balance of payments data highlight the intra-euro area fragmentation dynamics set off in March, on account of safe haven and flight-to-quality behaviour, whereby sizeable outflows from debt securities issued by more vulnerable euro area countries – Greece, Italy, Portugal and Spain – were recorded, while less vulnerable countries – i.e., Austria, Belgium, Finland, France, Germany and the Netherlands – experienced significant inflows – see Figure 3b. These debt inflows persisted and grew even

[3] This prominent role of investment funds is corroborated by Bergant et al. (2020), who find that, in response to the Eurosystem's APP, some sectors, of which households stand out, rebalanced their portfolios from domestic and other euro area debt securities towards foreign debt via their holdings of investment fund shares.

Figure 2: Portfolio investment flows by sector in 2020Q1 and 2020Q2 *(quarterly flows in % of euro area GDP)*



(a) Total assets

(b) Investment fund share assets

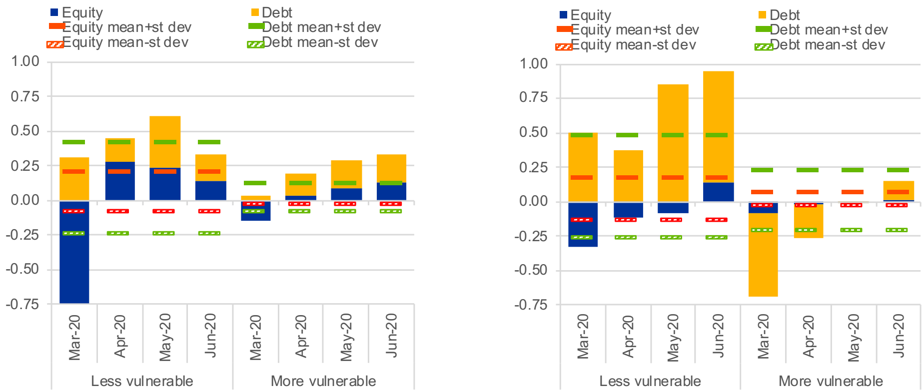
Source: ECB and Eurostat. "Domestic" refers to cases where the holder country is the same as the issuer country; "EA ex dom" refers to cases where the euro area holder country is not the same as the issuer country; "Non-EA" are all other countries. "MFI" and monetary financial institutions; "IF" are investment funds; "OFI" are other financial intermediaries; "IC" are insurance companies; "PF" are pension funds; "NFC" are non-financial corporations; "HH" are households. Conceptually, SHS and b.o.p. data are not fully compatible, to the extent that the latter do not include domestic flows.

larger in the second quarter, while more vulnerable countries' debt outflows declined in April, before turning into net inflows in June.

On the asset side, net sales of equity, which also encompass investment fund shares, were recorded for both groups of countries in March and were particularly large in the case of the less vulnerable countries – see Figure 3a. In both groups of countries, net purchases of equity were recorded since April, while net purchases of debt were recorded in March and throughout the second quarter. As regards the asset side, SHS data offer a more complete picture: Figures 4 and 5 – broken down by less and more vulnerable investor countries – show the role of euro area investment funds in the retrenchment-rebound dynamic in the first half of 2020, both in terms of the size of flows in investment fund shares and in that most of these flows were vis-à-vis euro area-based investment funds. However, there are clear differences between the two groups of countries, including the fact that less vulnerable countries resorted heavily to domestic investment funds, whereas more vulnerable countries transacted more with those based in other euro area countries, which also reflects the fact that more vulnerable countries have, on average, much lower shares of exposure to domestic investment funds (around a third in end-2019, whereas that of less vulnerable countries was closer to two thirds). Considering the aforementioned differences between the two groups, our analysis focuses on less and more vulnerable euro area countries, presenting their direct and indirect

Covid Economics 65, 20 January 2021: 153-187

Figure 3: Portfolio flows by country group *(monthly flows in % of euro area GDP)*



(a) Assets (b) Liabilities

Source: ECB and Eurostat. Notes: Averages and standard deviations calculated from January 2008 to June 2020. "Less vulnerable" countries are Austria, Belgium, Finland, France, Germany and the Netherlands; "more vulnerable" countries are Italy, Greece, Portugal and Spain.

exposures to securities at the end of 2019, before delving into the changes in the first two quarter of 2020, amid the COVID-19 shock.[4]

3 Estimation of indirect portfolio exposures

The method to look through investment fund holdings of euro area investors broadly follows Carvalho (2020). While in this study the focus was on the geography of holdings, we extend the focus to also include the issuer sectors of securities as well as their currency of denomination.

For that purpose, the Eurosystem's Securities Holdings Statistics (SHS) by sector are used, which contain highly granular security-level information on the holdings, valued at market prices, of debt securities, listed shares and investment fund shares of euro area investors, aggregated by institutional sectors.[5] It should be noted that this dataset represents a subset of euro area sectors' complete portfolio, to the extent that it does not include unlisted shares – in this sense, listed shares are taken as a proxy for the whole of the equity instrument class. [6]

The following sectors are considered and their holdings are accordingly aggregated at the euro area

[4]See Lane (2020a), Lane (2020b) and Lane (2020c).

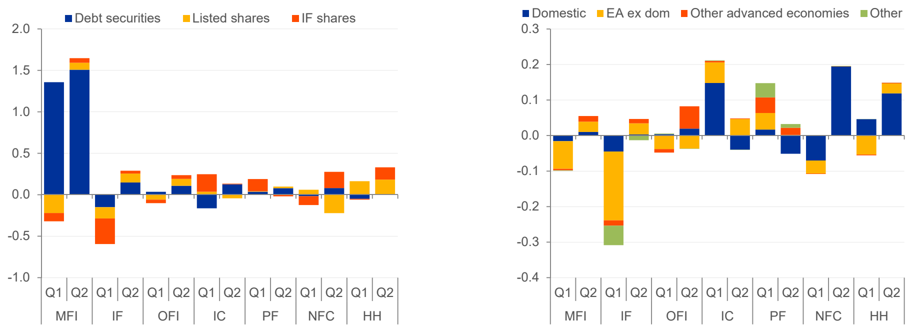
[5]Data for other EU countries that do not belong to the euro area are also available; however, these countries participate only on a voluntary basis, as opposed to euro area countries, for which the collection of this data is compulsory. In practical terms, the availability of data for the EU countries not belonging to the euro area is more restricted and, in general, of inferior quality.

[6]For more information on the SHS, see (ECB, 2015)

Covid Economics 65, 20 January 2021: 153-187

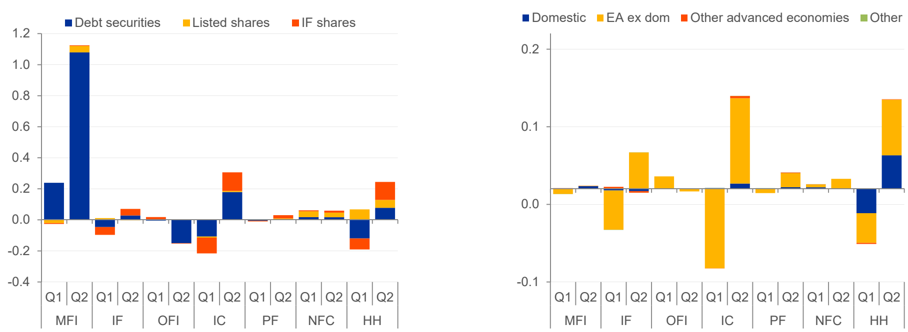
Covid Economics 65, 20 January 2021: 153-187

Figure 4: Portfolio investment flows by sector in 2020Q1 and 2020Q2 - less vulnerable countries *(quarterly flows in % of euro area GDP)*



(a) Total assets **(b)** Investment fund share assets

Figure 5: Portfolio investment flows by sector in 2020Q1 and 2020Q2 - more vulnerable countries *(quarterly flows in % of euro area GDP)*



(a) Total assets **(b)** Investment fund share assets

Source: ECB and Eurostat."Domestic" refers to cases where the holder country is the same as the issuer country; "EA ex dom" refers to cases where the euro area holder country is not the same as the issuer country; "Non-EA" are all other countries. "MFI" and monetary financial institutions; "IF" are investment funds; "OFI" are other financial intermediaries; "IC" are insurance companies; "PF" are pension funds; "NFC" are non-financial corporations; "HH" are households. Conceptually, SHS and b.o.p. data are not fully compatible, to the extent that the latter do not include domestic flows. "Less vulnerable" countries are Austria, Belgium, Finland, France, Germany and the Netherlands; "more vulnerable" countries are Italy, Greece, Portugal and Spain.

country-sector level: deposit taking corporations except central banks, which corresponds to banks (B); money market and non-money market investment funds (IF); insurance corporations (IC); pension funds

(PF); other financial institutions excluding financial vehicle corporations (OFI); non-financial corporations (NFC) and households (HH).

The first step towards "looking-through" is to compute the distribution of holdings by investment funds. Money-market and non-money market funds are pooled together into a single investment fund sector.[7]

This is done for the holdings of each instrument class and gives $w_{hc,IF,t}^{ic,a,is,c}$, i.e., the weight of asset class a, denominated in currency c, issued by issuer sector is of issuer country ic, in total investment fund (IF) holdings resident in euro area country hc, in period t. It should be noted that investment funds also hold investment fund shares (funds of funds) and therefore, in this step, these weights are computed for debt securities, listed shares, and also investment fund shares.

The second step is to compute the IF share holdings of each sector, in each euro area country and in each period of time. These are given by $h_{hc,hs,t}^{ic,IFS,IF,c}$, where hs stands for holder sector, the asset is investment fund shares (IFS), and the remaining super- and subscripts have the same interpretation as before. There is, however, one element which is unknown: the holdings of investment funds which are resident outside the euro area. In order to allocate these amounts, it is assumed that the distribution of these holdings corresponds to the average of that of Luxembourgian and Irish resident investment funds, which tend to be representative of global investment funds. Accordingly, that average is given by $\bar{w}_{hc,IF,t}^{ic,IFS,IF,c}$, where hc includes Luxembourg and Ireland.[8]

For this reason, one needs to separate the investment fund share holdings of a euro area country-sector vis-à-vis other euro area countries — $h_{hc,hs,t}^{ic=EA,IFS,IF,c}$, which we allocate on issuer-country by country basis — and non-euro area countries — $h_{hc,hs,t}^{ic\neq EA,IFS,IF,c} = \sum_{ic\notin EA} h_{hc,hs,t}^{ic,IFS,IF,c}$, which we allocate as single block of non-euro area issuer countries.

With all these elements, the estimated holdings of a given holder country hc of asset class a, denominated in currency c, issued by issuer sector is, vis-à-vis issuer country ic, in period t, is given by:

$$\hat{h}_{hc,hs,t,1}^{ic,a,is,c} = h_{hc,hs,t}^{ic=EA,IFS,IF,c} \times w_{hc,IF,t}^{ic,a,is,c} + h_{hc,hs,t}^{ic\neq EA,IFS,IF,c} \times \bar{w}_{hc,IF,t}^{ic,IFS,IF,c} \tag{1}$$

Due to the existence of funds of funds, this process is not able to distribute the whole investment fund

[7] In this pooled investment fund sector non-money market investment funds accounted for 92% of the overall holdings at end-2019.

[8] The potential bias introduced by this assumption is deemed not to be significant since the vast majority of investment fund shares held by euro area sectors is issued by investment funds resident in the euro area, for which the distribution is known. In fact, at end-2019, the average non-euro area share in total investment fund share holdings of euro area country-sectors was under 11%. For the countries which feature most prominently in this study – on the one hand, Austria, Belgium, Finland, France, Germany, the Netherlands and, on the other hand, Greece, Italy, Portugal and Spain – the share was even less than 6%.

Covid Economics 65, 20 January 2021: 153-187

Covid Economics 65, 20 January 2021: 153-187

shares holdings of a given country-sector pair in a single go. Therefore, a residual is attained, which is the difference between the observed holdings of investment fund shares per holding country-sector pair and the estimated figures, i.e. $\widehat{RES}^{ic,IFS,IF,c}_{hc,hs,t,1} = \sum_{ic} h^{ic,a,is,c}_{hc,hs,t} - \hat{h}^{ic,a,is,c}_{hc,hs,t,1}$. For this reason, the procedure is repeated, this time allocating only the residual:

$$\hat{h}^{ic,a,is,c}_{hc,hs,t,2} = \widehat{RES}^{ic=EA,IFS,IF,c}_{hc,hs,t,1} \times w^{ic,a,is,c}_{hc,IF,t} + \widehat{RES}^{ic\neq EA,IFS,IF,c}_{hc,hs,t,1} \times \bar{w}^{ic,IFS,IF,c}_{hc,IF,t} \tag{2}$$

This process is repeated until the residual is minimal — in practice, five rounds are sufficient to arrive at a residual which is about 0.1% of the original investment fund share holdings of each country-sector pair. Finally, the total estimated figures for debt instruments and listed shares are given by

$$\hat{h}^{ic,a,is,c}_{hc,hs,t} = h^{ic,a,is,c}_{hc,hs,t} + \hat{h}^{ic,a,is,c}_{hc,hs,t,1} + ... + \hat{h}^{ic,a,is,c}_{hc,hs,t,5} \tag{3}$$

We refer to $h^{ic,a,is,c}_{hc,hs,t}$ as the direct component of the investment of a given holding country-sector pair in asset class a (i.e., excluding the exposures held via investment funds) and $\hat{h}^{ic,a,is,c}_{hc,hs,t,1} + ... + \hat{h}^{ic,a,is,c}_{hc,hs,t,5}$ as the indirect component (i.e., the exposures held via investment funds).

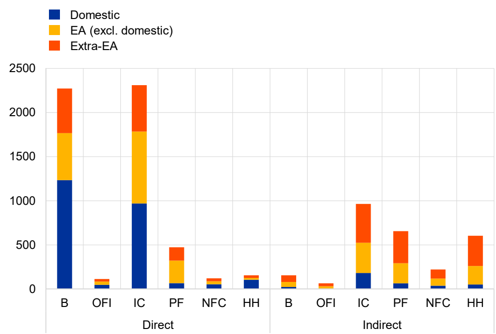
There are two caveats to our approach: the first is that the information in SHS data does not cover all investment fund assets, but only their portfolio investment holdings. For instance, SHS has no information on non-financial assets (such as real estate) held by euro area investment funds. Nevertheless, the instrument classes that are covered by the SHS make up for the bulk of the assets held by investment funds. In fact, according to the ECB's investment fund statistics, at the end of 2019, the share of debt securities, equities and investment fund shares in the aggregate balance sheet of euro area investment funds was slightly above 85%, with loans and deposits, non-financial assets and other assets accounting for the remainder.

The second caveat relates to the assumption that investors across euro area countries have the same preferences across issuing countries of investment fund shares, in terms of those funds' portfolio allocations. In practice this means that for investment fund shares issued by investment funds resident in Luxembourg, we need to assume that German investors hold fund shares issued by investment funds with the same asset composition as those held by French residents. The same applies for the different sectors, i.e., banks may invest predominantly in certain Irish investment funds, which are distinct from the choices of Irish investment funds made by pension funds. The concerns raised by these assumptions are partially alleviated by the findings in Monti and Felettigh (2008), who report that estimates of indirect holdings are not significantly affected by the assumption that they follow the overall distribution of Irish and Luxembourgian investment funds, compared to a more detailed estimation, where the investment strategy of the individual investment funds in which Italian residents invest is known.

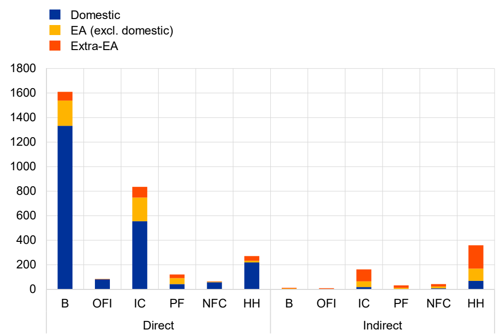
4 Descriptive evidence on direct and indirect holdings

Direct and indirect holdings of less and more vulnerable euro area countries to debt securities are displayed, respectively, in Figures 6a and 6b, and listed shares in Figures 7a and 7b.

Figure 6: Direct and indirect exposures of euro area investors to debt securities *(EUR bn, 2019-Q4)*



(a) Less vulnerable



(b) More vulnerable

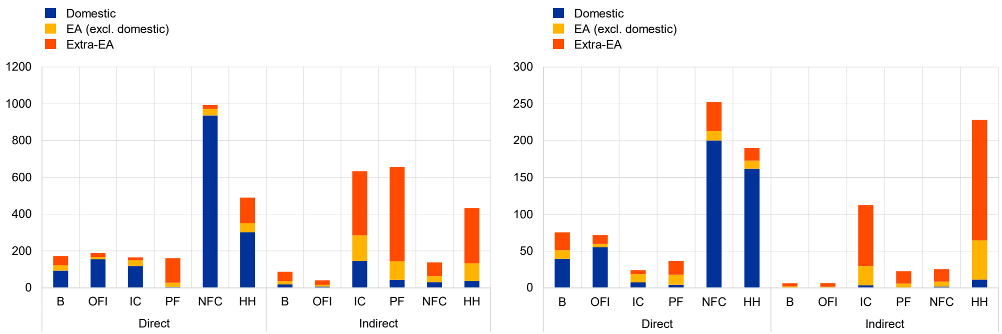
In general, indirect exposures tend to be more relevant in the case of listed shares than for debt securities. To the extent that home bias is lower in equity investment (see Fidora et al. (2007) on the reasons for lower home bias in equity as opposed to debt securities) it may render it more efficient to rely on investment funds to enter more distant markets/securities which carry higher information costs. Analogously, the lower information costs associated with bonds and in particular for sovereigns (which are present in fixed income markets only) may explain higher levels of direct exposures. Moreover, indirect exposures are much higher for foreign securities, in particular extra-euro area instruments. This implies, that traditional estimates of home-bias tend to be overestimated, as they do not take into account exposures via indirect holdings.

At the sectoral level, it is evident that those sectors with higher proportions of investment fund shares in their total portfolio have, mechanically, higher shares of indirect exposures. This is, for instance, the case for households which, in less vulnerable countries, have indirect exposures to debt securities which

Covid Economics 65, 20 January 2021: 153-187

Covid Economics 65, 20 January 2021: 153-187

Figure 7: Direct and indirect exposures of euro area investors to listed shares *(EUR bn, 2019-Q4)*



(a) Less vulnerable **(b)** More vulnerable

Source: Securities statistics, ECB, SHS and author's calculations. Notes: "Less vulnerable" countries are Austria, Belgium, Finland, France, Germany and the Netherlands; "more vulnerable" countries are Italy, Greece, Portugal and Spain. "Domestic" refers to cases where the holder country is the same as the issuer country; "EA excl. domestic" refers to cases where the euro area holder country is not the same as the issuer country; "Extra-EA" are all other countries. "B" are deposit-taking corporations; "OFI" are other financial intermediaries; "IC" are insurance companies; "PF" are pension funds; "NFC" are non-financial corporations; "HH" are households.

dwarf their direct exposures. The latter is in line with the relatively lower level of sophistication of this sector, which implies a higher propensity to resort to investment funds to attain international portfolio diversification. In the same vein, insurance companies, in both groups of countries, have larger indirect than direct exposures to listed shares, while pension funds from less vulnerable countries rely heavily on investment funds to obtain exposure to foreign debt securities and listed shares. On the contrary, banks have relatively small holdings of investment fund shares, which is why their indirect exposures are only a fraction of the direct holdings.

5 Empirical approach

We focus on the geography and issuer sector dimension of euro area portfolio investments. Accordingly, we use the following empirical approach, based on Galstyan and Lane (2013), but in a more dis-aggregated fashion:

$$\Delta ln(h^{ic,a,is}_{hc,hs}) = ln(h^{ic,a,is}_{hc,hs}) + \alpha^1_{hc} + \alpha^2_{ic,is} + \beta^1 \gamma_{hc,ic} + \beta^2 \eta_{hi} + \varepsilon_{hi,t} \tag{4}$$

where $\delta ln(h_{hc,hs}^{ic,a,is})$ is the difference in the log of the direct, indirect or estimated exposures, of holder country hc and holder sector hs to asset class a, of issuer country ic, issued by issuer sector is.

We consider two different periods to analyse the impact of the COVID epides: (i) from end-2019 to the end of the first quarter of 2020 ("shock period") and (ii) from the end of the first quarter to the end of the second quarter of 2020 ("rebound period"). We control for host country, source country and issuer sectors: holder country (α_{hc}^1) and issuer country/sector ($\alpha_{ic.is}^2$) dummies are included, in order to remove common trends and valuation effects, thereby ensuring that what remains is only the country-sector bilateral variation.[9]

In addition, three sets of controls are included in the regressions:

- $ln(h_{hc,hs}^{ic,a,is})$ is the log of the outstanding direct, indirect or estimated exposure at end-2019 and used to control for the pre-existing (i.e. before the COVID shock) level of a sector's investment.
- a set of gravity variables γ_{hi}, including the logarithms of bilateral distance and bilateral imports, as well as dummies for shared language (definitions and sources of these data are described in the appendix).
- a set of dummies η_{hi}, controlling for domestic exposures – i.e., whenever the holder country coincides with the issuer country – exposures vis-à-vis euro area less vulnerable countries – i.e, whenever the issuer countries are Austria, Belgium, Finland, France, Germany and the Netherlands, and excluding domestic exposures – exposures vis-à-vis euro area more vulnerable countries – Greece, Italy, Portugal and Spain – and exposures to the remaining non-euro area countries. Due to the lack of gravity variables, the portfolio exposures to some territories were reclassified and included within larger ones – the list of reclassifications is also provided in the appendix – and, for the same reason, exposures to debt issued by international organisations was excluded. In most cases, these exposures are relatively small, except for the cases of non-EU European financial centres (such as Guernsey, Jersey and Lichtenstein), territories vis-à-vis which euro area residents tend to have larger exposures.

6 Results: portfolio shifts during the COVID-19 pandemic

This section present the results obtained for the COVID period, first for debt securities and then listed shares. In both cases, it starts by looking at the COVID shock period for the whole set of the euro

[9]In order to focus only on non-trivial holdings, as well as to avoid potential bias, exposures smaller than €1 million were excluded.

Covid Economics 65, 20 January 2021: 153-187

area countries, investors resident in less vulnerable countries and those resident in the more vulnerable countries, followed by an equivalent analysis for the "rebound" period.

6.1 Debt securities: shock period

6.1.1 All euro area countries

The estimation results for debt securities are presented across the different holding sectors, i.e. banks, OFI, IC, PF, NFC and HH and show the drivers of shifts in direct, indirect and total exposures during the first quarter of 2020. While the main focus is on the results for sector-geographical dummies (i.e. issuance by banks, governments and non-financial corporations, respectively, resident in the domestic economy, euro area less vulnerable and more vulnerable countries as well as outside the euro area, respectively), we first present the findings for the control variables outlined in equation (4).

Table 1 shows the results for all euro area countries for shifts in holdings of debt securities. In line with Galstyan and Lane (2013) and Mehl et al. (2019) for the global financial crisis, there is a significant "reversion to the mean" effect across almost all exposures and investor sectors, implying that those positions that were larger at the end of 2019, were most reduced during the first quarter of 2020; moreover, there is evidence that distance matters, in particular for direct holdings: banks, OFI, IC and HH reduced positions more in remote locations. Interestingly, shifts in indirect positions were also negatively affected by distance in the case of banks, OFI and PF.

As regards geography and issuer sectors, the striking general feature emerges that rebalancing patterns were more significant once shifts in indirect exposures are included in the analysis. The strongest rebalancing in 2020Q1 affected sovereign debt securities: all euro area sectors rebalanced towards domestic sovereign debt which, in the case of banks and HH, was driven by direct exposures, while for OFI and NFC this pattern only emerges once indirect exposures are considered. Similarly, all euro area sectors rebalanced away from sovereign debt issued outside the euro area, which was driven by indirect holdings for banks, OFI and NFC, while for all other sectors indirect exposures were also significant. Hence, the re-orientation towards domestic sovereigns at the expense of extra-euro area sovereign debt was a remarkable feature of the rebalancing observed during the first quarter of 2020, and goes even beyond the patterns predicted by gravity. Plausibly, this pattern was driven by flight-to-safety considerations (Lane (2020a)), home and familiarity biases as well as anticipation of valuation gains from an increased size of asset purchases programmes by the ECB (as observed before the APP, see Bergant et al. (2020)). Furthermore, Broner et al. (2014) argue that investors may have an incentive to buy domestic debt to

Covid Economics 65, 20 January 2021: 153-187

the extent they are positively discriminated vis-à-vis foreign investors, since they have a lower probability of being negatively affected in a default episode; domestic investors may also buy domestic debt due to moral suasion.[10]

Less clear-cut pictures are found for issuance by banks and NFC: OFI and HH rebalanced away from debt securities issued by banks resident outside the euro, and consistently so across direct and indirect holdings. There is some evidence of a rebalancing into debt securities issued by NFC, in particular for IC (into domestic and other euro area debt, driven by indirect exposures) and for NFC (across all geographic locations and partly due to indirect exposures).

Covid Economics 65, 20 January 2021: 153-187

[10]See also Broner et al. (2010) and Altavilla et al. (2017).

Table 1: All euro area countries - debt securities - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.03***	-0.07***	-0.03***	-0.17***	-0.05***	-0.16***	-0.03***	-0.02***	-0.03***	-0.05***	-0.03***	-0.04***	-0.07***	-0.02*	-0.06***	-0.03***	-0.01	-0.02**
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Distance	-0.05**	-0.07***	-0.04*	-0.22***	-0.02*	-0.14***	-0.02	-0.01*	-0.01*	-0.02	-0.01*	-0.02	-0.04	-0.00	-0.04*	-0.05**	-0.01	-0.02***
	(0.02)	(0.02)	(0.02)	(0.05)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.04)	(0.00)	(0.02)	(0.04)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Imports	0.00	-0.00	0.00	-0.00	-0.01**	0.01	0.00	0.00	-0.00	0.03	-0.00	0.01	0.01	0.00	0.00	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.01)	(0.02)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Com. language	-0.03	-0.16***	-0.03	-0.11	0.10***	-0.03	0.02	0.01	0.02**	-0.12***	0.02**	0.01	0.14	-0.01	0.02	-0.02	0.01	-0.00
	(0.05)	(0.04)	(0.05)	(0.10)	(0.02)	(0.06)	(0.02)	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.10)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)
B-DOM	0.60	0.11	0.49*	0.58*	-0.24	-0.06	-0.01	0.37**	-0.41***	0.17	0.22	-0.11	0.22	-0.40	-0.53**	0.05	0.17	0.03
	(0.42)	(0.11)	(0.28)	(0.32)	(0.38)	(0.14)	(0.06)	(0.18)	(0.12)	(0.10)	(0.34)	(0.14)	(0.36)	(0.51)	(0.23)	(0.14)	(0.24)	(0.15)
B-LESSVUL	0.04	0.09	-0.04	-1.43**	-0.73*	-1.24**	-0.11	0.18	-0.23**	-0.17	0.02	-0.22	-0.09	-0.58	-0.17	-0.26	-0.07	-0.16
	(0.12)	(0.10)	(0.11)	(0.64)	(0.38)	(0.52)	(0.07)	(0.18)	(0.12)	(0.10)	(0.33)	(0.16)	(0.26)	(0.51)	(0.24)	(0.16)	(0.22)	(0.14)
B-MOREVUL	0.02	0.19***	-0.04	-1.16***	-0.66*	-1.93**	-0.02	0.17	-0.18	-0.11	-0.01	-0.12	0.01	-0.65	-0.15	-0.35**	-0.04	-0.19
	(0.18)	(0.06)	(0.13)	(0.34)	(0.38)	(0.97)	(0.09)	(0.18)	(0.11)	(0.13)	(0.33)	(0.15)	(0.25)	(0.51)	(0.24)	(0.14)	(0.22)	(0.14)
B-NEA	-0.56	-0.08	-0.48*	-1.19***	-0.45***	-0.38***	-0.00	-0.17***	0.24***	0.02	-0.18***	0.07	0.27	-0.15***	0.60***	-0.24*	-0.23***	-0.23***
	(0.40)	(0.07)	(0.26)	(0.29)	(0.06)	(0.11)	(0.06)	(0.04)	(0.04)	(0.18)	(0.05)	(0.10)	(0.32)	(0.04)	(0.07)	(0.13)	(0.07)	(0.07)
GG-DOM	0.52*	0.19	0.92***	-0.09	0.35*	0.29	0.48***	0.35***	0.36***	0.56***	0.37***	0.39***	0.46	0.61***	0.38***	0.51***	0.28**	0.31***
	(0.29)	(0.23)	(0.23)	(0.38)	(0.21)	(0.30)	(0.07)	(0.06)	(0.06)	(0.15)	(0.06)	(0.12)	(0.29)	(0.16)	(0.09)	(0.13)	(0.11)	(0.10)
GG-LESSVUL	0.24	-0.17	0.19	-0.35	-0.44**	-0.06	0.06	-0.01	0.10	-0.31*	-0.11	-0.14	0.54	0.12	0.14	0.11	-0.18	-0.10
	(0.19)	(0.23)	(0.18)	(0.55)	(0.21)	(0.48)	(0.10)	(0.08)	(0.09)	(0.18)	(0.12)	(0.09)	(0.56)	(0.19)	(0.35)	(0.09)	(0.14)	(0.11)
GG-MOREVUL	0.27*	-0.31	0.20	-0.75**	-0.47**	-0.47*	-0.04	-0.13	0.01	-0.10	-0.14	-0.05	-0.31	-0.02	-0.26**	0.08	-0.20	-0.03
	(0.16)	(0.24)	(0.16)	(0.37)	(0.24)	(0.26)	(0.07)	(0.10)	(0.05)	(0.20)	(0.09)	(0.14)	(0.23)	(0.18)	(0.13)	(0.11)	(0.14)	(0.07)
GG-NEA	-0.33	-0.43***	-0.74***	-0.28	-0.69***	-0.36*	-0.41***	-0.31***	-0.29***	-0.37***	-0.35***	-0.34***	-0.14	-0.35***	-0.31***	-0.41***	-0.36***	-0.35***
	(0.24)	(0.01)	(0.17)	(0.29)	(0.05)	(0.22)	(0.05)	(0.04)	(0.04)	(0.08)	(0.04)	(0.08)	(0.22)	(0.03)	(0.04)	(0.10)	(0.07)	(0.07)
NFC-MOREVUL	0.60***	0.02	0.47**	-0.28	0.24	-0.01	0.22	-0.06	0.14**	0.74***	0.09*	0.06	0.88***	0.21	0.78***	-0.09	0.17*	-0.11
	(0.23)	(0.13)	(0.22)	(1.96)	(0.14)	(0.60)	(0.14)	(0.05)	(0.07)	(0.22)	(0.05)	(0.14)	(0.31)	(0.13)	(0.24)	(0.24)	(0.10)	(0.26)
NFC-LESSVUL	0.21	0.15	0.20	-0.74	-0.20	-0.27	0.12	-0.12***	0.14**	-0.08	0.02	-0.08	0.28	0.09	0.64***	-0.56**	0.08	-0.20
	(0.18)	(0.13)	(0.17)	(0.46)	(0.15)	(0.36)	(0.09)	(0.04)	(0.07)	(0.17)	(0.05)	(0.11)	(0.23)	(0.14)	(0.24)	(0.24)	(0.08)	(0.26)
NFC-MOREVUL	0.03	0.09	0.10	-0.37	-0.10	-0.07	0.13	-0.20***	0.17**	-0.23*	-0.03	-0.11	0.86**	0.02	0.94***	-0.47**	0.06	-0.15
	(0.22)	(0.11)	(0.20)	(0.47)	(0.16)	(0.35)	(0.09)	(0.04)	(0.07)	(0.12)	(0.05)	(0.11)	(0.34)	(0.14)	(0.27)	(0.23)	(0.09)	(0.25)
NFC-NEA	-0.39***	0.07	-0.24*	-0.24	-0.33***	-0.00	-0.02	-0.01	0.04	-0.62***	-0.01	-0.05	-0.10	-0.04	0.07	-0.28***	-0.07	-0.06
	(0.15)	(0.09)	(0.14)	(1.91)	(0.05)	(0.49)	(0.11)	(0.04)	(0.04)	(0.05)	(0.04)	(0.08)	(0.22)	(0.03)	(0.06)	(0.07)	(0.07)	(0.07)
Constant	0.52**	0.42***	0.37	2.28***	0.49***	1.54***	0.24*	-0.05	-0.00	0.02	-0.05	0.03	0.33	-0.11	0.20	0.34	-0.02	0.13
	(0.24)	(0.15)	(0.22)	(0.54)	(0.08)	(0.30)	(0.14)	(0.06)	(0.08)	(0.51)	(0.05)	(0.21)	(0.50)	(0.09)	(0.17)	(0.26)	(0.08)	(0.10)
Observations	1,694	1,280	2,118	999	1,858	2,084	2,412	3,484	3,838	1,270	2,470	2,641	802	1,953	2,107	1,171	3,183	3,312
R-squared	0.31	0.89	0.34	0.43	0.93	0.45	0.37	0.87	0.61	0.40	0.88	0.68	0.33	0.84	0.49	0.43	0.87	0.78

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

Covid Economics 65, 20 January 2021: 153-187

CEPR PRESS

Covid Economics 65, 20 January 2021: 153-187

6.1.2 Less vulnerable euro area countries

Next, we zoom in on portfolio shifts in debt securities of the less vulnerable euro area countries (Table 2). Increased total exposures to domestic sovereign debt is found for most sectors, with the exception of banks and NFC. For the other sectors this is driven by indirect exposures, while for HH also direct exposures contribute. Moreover, there was a shift towards sovereign debt issued in other euro area countries (both vulnerable and less vulnerable countries) for all sectors, except for banks. In the cases of OFI, IC, PF and NFC this was exclusively due to indirect holdings. A rebalancing away from non-euro area sovereign debt is found for most sectors (except for NFC and HH) and driven by indirect exposures for banks and IC.

Investors from less vulnerable countries (except for banks) also increased significantly their exposures to domestic and less vulnerable euro area countries' banks. Again this is driven by indirect holdings. For debt securities issued by NFCs, HH and IC increased their total exposures significantly across all geographical dimensions (mainly due to indirect holdings), while banks, NFC and OFI concentrated their expansion on domestic and other NFC debt of less vulnerable countries.

6.1.3 More vulnerable euro area countries

Compared to the less vulnerable countries, there is generally less evidence of significant rebalancing in debt securities during 2020Q1 for the more vulnerable countries, as can be seen in Table 3. Shifts in exposures were mainly focused on sovereign debt. Most strikingly, exposures to domestic sovereign debt increased across the board (largely via indirect holdings), in an environment of a general sell-off of these securities by foreign investors (see Figure 3b). Exposures to other euro area sovereign debt increased only for IC (directly towards debt issued by less vulnerable countries) and HH (via indirect holdings both for less and more vulnerable countries). As for the less vulnerable countries, there was a reduction in exposures to sovereign debt issued outside the euro area (mainly via indirect for banks, IC, PF and NFC as well as directly and indirectly for HH).

As regards holdings of debt issued by banks, the clearest results are found for HH: exposures to domestic and less vulnerable bank debt increased, while those toward extra-euro area banks declined (all driven by indirect holdings). Among debt securities issued by NFC, mainly PF were active, by rebalancing into domestic and less vulnerable securities, while reducing exposures to extra-euro area NFCs (the latter entirely due to indirect exposures).

Table 2: Less vulnerable euro area countries - debt securities - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h_{hc,hs}^{ic,a,is})$	-0.01	0.00	0.00	-0.30***	-0.02**	-0.33***	-0.05***	-0.03***	-0.04***	-0.03	-0.03**	-0.05***	-0.03	-0.02**	-0.10**	-0.02*	-0.03***	-0.04***
	(0.02)	(0.02)	(0.02)	(0.06)	(0.01)	(0.04)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)
Distance	-0.01	0.00	0.01	-0.38***	-0.01	-0.29***	-0.06***	-0.00	-0.02	-0.09*	-0.01	-0.08***	-0.09	0.03*	-0.05	-0.05	-0.00	-0.01
	(0.05)	(0.02)	(0.04)	(0.12)	(0.01)	(0.06)	(0.02)	(0.01)	(0.01)	(0.05)	(0.01)	(0.03)	(0.09)	(0.02)	(0.04)	(0.03)	(0.02)	(0.02)
Imports	0.01	-0.01	0.00	0.03	-0.01*	0.00	-0.00	-0.00	-0.00	0.01	-0.00	-0.00	-0.01	0.00	0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.04)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.03)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Com. language	0.04	-0.02	0.03	-0.18	0.03*	-0.20**	-0.01	0.03	0.05**	-0.21***	0.02	0.05*	0.18	0.02	0.07	0.07*	0.03*	0.04**
	(0.13)	(0.03)	(0.10)	(0.25)	(0.02)	(0.10)	(0.04)	(0.02)	(0.02)	(0.07)	(0.02)	(0.03)	(0.19)	(0.02)	(0.06)	(0.04)	(0.02)	(0.02)
B-DOM	0.81	0.35**	0.74	0.81	0.47***	1.34**	-0.05	0.40***	0.28***	-0.53**	0.37***	0.23	0.24	0.48***	0.72***	0.18	0.44***	0.39***
	(0.68)	(0.16)	(0.50)	(0.82)	(0.08)	(0.54)	(0.14)	(0.08)	(0.09)	(0.23)	(0.07)	(0.16)	(0.49)	(0.08)	(0.20)	(0.18)	(0.09)	(0.10)
B-LESSVUL	0.35	0.23*	0.01	-0.85	-0.12*	0.73**	0.02	0.33***	0.41***	-0.31*	0.22***	0.08	-0.00	0.26***	0.52***	-0.17	0.34***	0.30***
	(0.32)	(0.13)	(0.22)	(0.56)	(0.07)	(0.32)	(0.14)	(0.10)	(0.12)	(0.17)	(0.09)	(0.14)	(0.37)	(0.08)	(0.20)	(0.12)	(0.12)	(0.10)
B-MOREVUL	-0.25	0.19***	-0.28	-1.31**	-0.23***	-0.21	-0.27	0.12	0.14	-0.63*	-0.04	-0.08	-0.23	0.13**	0.12	-0.02	0.18**	0.21***
	(0.48)	(0.07)	(0.17)	(0.52)	(0.03)	(0.27)	(0.25)	(0.11)	(0.12)	(0.33)	(0.17)	(0.16)	(0.30)	(0.05)	(0.13)	(0.08)	(0.08)	(0.07)
B-NEA	-0.37	-0.22**	-0.65	-0.62	-0.59***	0.20	0.06	-0.05	0.16*	0.24***	-0.15**	-0.11	-0.04	-0.12***	0.14	-0.23**	-0.04	-0.01
	(0.72)	(0.10)	(0.47)	(0.40)	(0.05)	(0.15)	(0.07)	(0.06)	(0.09)	(0.08)	(0.07)	(0.10)	(0.33)	(0.04)	(0.14)	(0.12)	(0.07)	(0.06)
GG-DOM	-0.01	0.59***	0.50*	0.97	0.55***	0.78*	0.31	0.60***	0.57***	0.31	0.68***	0.56***	0.22	0.77***	0.77	0.71***	0.65***	0.67***
	(0.50)	(0.13)	(0.28)	(0.83)	(0.09)	(0.43)	(0.31)	(0.11)	(0.12)	(0.25)	(0.14)	(0.17)	(1.27)	(0.13)	(0.67)	(0.21)	(0.15)	(0.12)
GG-LESSVUL	0.35	0.24**	0.02	-1.22**	-0.12**	0.54**	0.03	0.31***	0.41***	-0.28	0.18**	0.07	0.39	0.25***	0.57	0.34*	0.34***	0.37***
	(0.33)	(0.10)	(0.21)	(0.62)	(0.06)	(0.27)	(0.13)	(0.09)	(0.12)	(0.22)	(0.09)	(0.15)	(1.03)	(0.08)	(0.48)	(0.20)	(0.12)	(0.09)
GG-MOREVUL	0.86**	0.27**	0.51*	-0.76	-0.10	0.67**	0.19	0.33***	0.47***	-0.19	0.17**	0.15	0.03	0.24***	0.51***	0.36	0.40***	0.41***
	(0.38)	(0.12)	(0.30)	(0.99)	(0.06)	(0.33)	(0.12)	(0.08)	(0.11)	(0.13)	(0.07)	(0.11)	(0.30)	(0.06)	(0.17)	(0.23)	(0.10)	(0.08)
GG-NEA	0.24	-0.46***	-0.61***	-1.73***	-0.64***	-0.25**	-0.35	-0.21***	-0.18***	-0.49***	-0.34***	-0.34***	-0.38	-0.33***	-0.27***	-0.37***	-0.18**	-0.16**
	(0.51)	(0.02)	(0.15)	(0.40)	(0.03)	(0.13)	(0.25)	(0.07)	(0.07)	(0.07)	(0.06)	(0.10)	(0.27)	(0.04)	(0.07)	(0.08)	(0.07)	(0.07)
NFC-DOM	0.72***	0.31*	0.47*	0.03	0.29***	0.24	-0.16	0.33***	0.36***	0.13	0.27***	0.06	-0.07	0.46***	0.60***	0.39**	0.32***	0.37***
	(0.27)	(0.18)	(0.25)	(1.13)	(0.09)	(0.88)	(0.16)	(0.08)	(0.09)	(0.30)	(0.07)	(0.17)	(0.55)	(0.08)	(0.19)	(0.16)	(0.10)	(0.09)
NFC-LESSVUL	0.41	0.27**	0.10	-0.34	-0.13**	0.87***	0.09	0.33***	0.46***	-0.44***	0.21***	0.04	-0.12	0.27***	0.44***	-0.02	0.33***	0.33***
	(0.35)	(0.11)	(0.20)	(0.61)	(0.06)	(0.21)	(0.12)	(0.09)	(0.11)	(0.17)	(0.08)	(0.14)	(0.37)	(0.08)	(0.16)	(0.10)	(0.11)	(0.08)
NFC-MOREVUL	-0.00	0.15*	-0.25	-0.65	-0.27***	0.14	-0.27**	0.17**	0.20**	-0.56**	0.03	-0.07	-0.33	0.10*	0.08	-0.16	0.20**	0.21***
	(0.47)	(0.08)	(0.28)	(0.44)	(0.04)	(0.17)	(0.11)	(0.07)	(0.09)	(0.24)	(0.06)	(0.11)	(0.37)	(0.06)	(0.15)	(0.11)	(0.09)	(0.07)
NFC-NEA	-0.14	-0.11*	-0.22**	0.05	-0.44***	0.93	0.25***	0.07	0.15**	-0.71***	-0.01	-0.07	0.04	-0.05*	0.13	-0.25***	0.08	0.08*
	(0.33)	(0.06)	(0.11)	(1.00)	(0.03)	(0.79)	(0.04)	(0.06)	(0.07)	(0.09)	(0.06)	(0.10)	(0.34)	(0.03)	(0.10)	(0.06)	(0.05)	(0.05)
Constant	-0.36	-0.08	-0.16	3.58***	0.42***	2.88***	0.60***	-0.16	-0.05	0.89*	-0.06	0.60**	0.83	-0.39***	0.32	0.27	-0.18	-0.09
	(0.66)	(0.18)	(0.36)	(1.29)	(0.09)	(0.59)	(0.19)	(0.13)	(0.13)	(0.48)	(0.09)	(0.27)	(1.04)	(0.14)	(0.33)	(0.33)	(0.15)	(0.15)
Observations	705	762	974	433	1,114	1,141	967	1,651	1,669	523	1,148	1,153	475	1,242	1,260	546	1,505	1,521
R-squared	0.41	0.88	0.49	0.57	0.97	0.65	0.47	0.84	0.75	0.72	0.89	0.84	0.41	0.85	0.52	0.64	0.90	0.87

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

Covid Economics 65, 20 January 2021: 153-187

Table 3: More vulnerable euro area countries - debt securities - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.04*	-0.08***	-0.03*	-0.24**	0.04*	-0.11*	-0.01	-0.02	-0.01	-0.04*	-0.02*	-0.04**	-0.19*	0.02	-0.01	-0.04*	0.00	-0.05**
	(0.02)	(0.02)	(0.02)	(0.09)	(0.02)	(0.06)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.10)	(0.01)	(0.03)	(0.02)	(0.01)	(0.02)
Distance	-0.08	-0.07**	-0.12**	-0.18	0.05	-0.22	0.02	-0.03***	0.02	-0.05	-0.01	-0.01	-0.16	-0.01	-0.07	-0.05	-0.00	-0.02*
	(0.08)	(0.03)	(0.06)	(0.19)	(0.03)	(0.19)	(0.04)	(0.01)	(0.03)	(0.05)	(0.01)	(0.03)	(0.28)	(0.02)	(0.05)	(0.06)	(0.01)	(0.01)
Imports	-0.00	-0.02**	-0.01	0.02	-0.00	0.06	0.00	0.00	-0.00	0.00	0.00	0.01	-0.05	0.01	-0.01	0.00	-0.00	-0.00
	(0.02)	(0.01)	(0.01)	(0.06)	(0.01)	(0.05)	(0.02)	(0.00)	(0.01)	(0.02)	(0.00)	(0.01)	(0.05)	(0.00)	(0.01)	(0.03)	(0.00)	(0.00)
Com. language	-0.21	0.00	-0.04	-0.23	0.03	0.31	-0.01	-0.00	-0.00	-0.10	0.00	-0.07*	-0.08	-0.01	-0.03	-0.03	-0.02	-0.04**
	(0.20)	(0.05)	(0.16)	(0.48)	(0.05)	(0.43)	(0.05)	(0.01)	(0.02)	(0.12)	(0.01)	(0.04)	(0.17)	(0.01)	(0.03)	(0.11)	(0.01)	(0.02)
B-DOM	0.00	0.03	0.13	0.19	0.22**	0.81	-0.08	0.22***	-0.17	-0.21	0.23***	-0.14	-0.13	0.21**	-0.72***	0.24	0.16***	0.35**
	(0.27)	(0.15)	(0.34)	(1.04)	(0.11)	(1.03)	(0.23)	(0.04)	(0.15)	(0.23)	(0.07)	(0.17)	(0.58)	(0.09)	(0.20)	(0.24)	(0.05)	(0.15)
B-LESSVUL	-0.05	-0.00	-0.05	0.00	0.33***	-0.43	0.27***	0.13	0.09	0.13	0.13***	0.17***	-0.28	0.19***	0.03	0.18	0.01	0.27**
	(0.16)	(0.08)	(0.11)	(0.63)	(0.12)	(0.33)	(0.09)	(0.09)	(0.10)	(0.09)	(0.02)	(0.06)	(0.62)	(0.04)	(0.11)	(0.28)	(0.07)	(0.13)
B-MOREVUL	-0.01	-0.38***	-0.07	-0.34	0.22	-1.23	0.31	-0.01	0.08	-0.17	-0.06	-0.14	-0.81	0.10	-0.17	-0.28	-0.13***	0.01
	(0.30)	(0.12)	(0.26)	(0.65)	(0.15)	(1.01)	(0.29)	(0.05)	(0.20)	(0.17)	(0.06)	(0.13)	(0.72)	(0.08)	(0.20)	(0.23)	(0.05)	(0.08)
B-NEA	-0.05	-0.59***	-0.44	0.41	0.01	-0.33	0.40**	-0.28***	0.21**	0.01	-0.28***	0.03	-0.53	-0.08	0.46	-0.18	-0.33***	-0.28***
	(0.18)	(0.14)	(0.28)	(0.61)	(0.11)	(0.33)	(0.20)	(0.07)	(0.11)	(0.21)	(0.04)	(0.15)	(0.82)	(0.06)	(0.30)	(0.35)	(0.05)	(0.04)
GG-DOM	0.55	0.53***	0.32	0.52	0.78***	0.14	0.24	0.66***	0.60***	0.33	0.77***	0.75***	0.52	0.71***	0.65**	0.68*	0.71***	0.72***
	(0.50)	(0.17)	(0.35)	(1.11)	(0.25)	(0.82)	(0.29)	(0.19)	(0.17)	(0.41)	(0.11)	(0.25)	(1.13)	(0.15)	(0.31)	(0.36)	(0.17)	(0.15)
GG-LESSVUL	-0.03	-0.06	-0.08	-0.44	0.37***	-1.04*	0.42***	0.11	0.21**	0.02	0.14***	0.02	-0.50	0.25***	0.09	0.23	0.05	0.28**
	(0.20)	(0.10)	(0.16)	(0.84)	(0.13)	(0.58)	(0.11)	(0.07)	(0.10)	(0.22)	(0.03)	(0.17)	(0.59)	(0.07)	(0.14)	(0.27)	(0.06)	(0.11)
GG-MOREVUL	0.18	0.09	0.10	-0.05	0.42*	-0.80	0.31***	0.10	0.11	0.21	0.16**	0.19	-0.27	0.31***	0.11	0.11	0.06	0.17
	(0.23)	(0.12)	(0.21)	(0.78)	(0.22)	(0.80)	(0.12)	(0.15)	(0.12)	(0.24)	(0.07)	(0.17)	(0.59)	(0.10)	(0.18)	(0.17)	(0.11)	(0.12)
GG-NEA	-0.47	-0.46***	-0.46*	-0.28	-0.21	-0.64	0.17*	-0.43***	-0.45***	-0.16	-0.44***	-0.54***	-0.21	-0.21***	-0.46**	-0.44*	-0.51***	-0.44***
	(0.43)	(0.12)	(0.26)	(0.44)	(0.13)	(0.50)	(0.09)	(0.04)	(0.05)	(0.16)	(0.05)	(0.10)	(0.42)	(0.08)	(0.22)	(0.23)	(0.02)	(0.04)
NFC-DOM	0.06	-0.21	-0.23	-0.49	0.08	-0.06	0.03	0.09	0.14	-0.36	0.05	0.14	0.13	-0.00	-0.01	0.33	0.01	0.02
	(0.27)	(0.13)	(0.22)	(1.08)	(0.15)	(0.70)	(0.28)	(0.07)	(0.18)	(0.25)	(0.05)	(0.13)	(1.20)	(0.10)	(0.43)	(0.44)	(0.09)	(0.13)
NFC-LESSVUL	0.33	-0.15*	0.16	0.77	0.27*	-0.04	0.36***	0.04	0.15	0.19*	0.05	0.15***	-0.37	0.12**	-0.02	0.09	-0.06	0.18
	(0.30)	(0.09)	(0.20)	(0.74)	(0.14)	(0.55)	(0.11)	(0.08)	(0.11)	(0.10)	(0.03)	(0.05)	(0.61)	(0.05)	(0.11)	(0.15)	(0.06)	(0.12)
NFC-MOREVUL	-0.22	-0.32**	-0.26	-0.30	0.15	-0.86	0.30**	-0.06	0.08	-0.09	-0.11**	-0.08	0.68	0.06	0.57	-0.25	-0.19***	-0.09
	(0.24)	(0.15)	(0.17)	(0.80)	(0.17)	(0.53)	(0.12)	(0.07)	(0.09)	(0.14)	(0.05)	(0.08)	(1.08)	(0.08)	(0.54)	(0.20)	(0.05)	(0.08)
NFC-NEA	-0.11	-0.18**	-0.04	0.80	0.14	-0.34	0.29***	-0.07	-0.10**	0.29***	-0.09**	-0.14	-0.10	0.13*	-0.03	-0.37*	-0.16***	-0.01
	(0.16)	(0.09)	(0.13)	(0.61)	(0.12)	(0.33)	(0.07)	(0.05)	(0.04)	(0.05)	(0.05)	(0.09)	(0.34)	(0.07)	(0.15)	(0.20)	(0.04)	(0.06)
Constant	0.92	0.70**	1.17*	1.82	-0.30	2.06	-0.45	0.17*	-0.22	0.36	-0.01	0.06	2.57	-0.28	0.68	0.34	0.08	0.31**
	(0.78)	(0.29)	(0.60)	(1.82)	(0.32)	(1.75)	(0.49)	(0.09)	(0.28)	(0.52)	(0.11)	(0.29)	(2.45)	(0.20)	(0.62)	(0.76)	(0.12)	(0.13)
Observations	413	428	548	179	339	376	519	800	845	388	624	666	191	591	608	359	972	990
R-squared	0.62	0.86	0.59	0.71	0.93	0.57	0.67	0.97	0.72	0.63	0.98	0.83	0.64	0.93	0.77	0.62	0.95	0.91

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

6.2 Debt securities: rebound period

There are stark differences between the shock and the rebound period, when financial conditions eased markedly compared to the initial COVID-19 shock of the first quarter. Specifically, there is no longer a common discernible pattern indicative of increased home bias or flight-to-safety (Table 4). Having said this, it is striking that euro area investors exposures to euro area sovereign debt remained roughly unchanged during this quarter, in spite of the large-scale net purchases by the ECB, possibly induced by the large volume of net issuance of euro area sovereign debt to finance governments' crisis response.

At the same time, HH and IC increased their exposure to extra-euro area sovereign debt – partly via indirect holdings – and thereby reversing some of the rebalancing undertaken in the first quarter, and also to extra-euro area debt securities issued by NFC. A common feature across sectors is, to some extent, a decreasing exposure to debt securities issued by non-euro area resident banks, largely via direct holdings.

In less vulnerable countries (Table 5) the most significant shifts are visible for OFI and HH, which increase their non-domestic exposures to bank and NFC debt, as well as sovereign, in the case of OFI. Importantly, these increases are due to indirect holdings. As regards more vulnerable countries, mostly IC, and to a lesser extent HH, display the same rebound patterns as for the euro area as whole, namely towards non-euro area sovereign and NFC debt (Table 6).

All in all, this evidence is in line with the significant improvement observed in financial markets in the second quarter of 2020, against the backdrop of the comprehensive measures undertaken by fiscal and monetary authorities. Regarding the latter, the deployment of the ECB's Pandemic Emergency Purchase Programme (PEPP) proved instrumental in restoring the functioning of markets and considerably reducing heightened risk aversion (Lane (2020b)).

Covid Economics 65, 20 January 2021: 153-187

Table 4: All euro area countries - debt securities - 2020Q1-2020Q2

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.06***	-0.05***	-0.06***	-0.10***	-0.11***	-0.10***	-0.03***	-0.04***	-0.03***	-0.05***	-0.02**	-0.06***	-0.06***	-0.05***	-0.03***	-0.02**	-0.05***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Distance	-0.08***	0.00	-0.07**	-0.04	-0.02	-0.05**	-0.02	-0.01	-0.03**	-0.01	0.01**	-0.01	-0.03	0.00	-0.02	-0.03	0.00	-0.01
	(0.03)	(0.01)	(0.03)	(0.04)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Imports	-0.00	0.01*	-0.00	-0.01	0.01***	0.00	-0.00	0.01***	0.00	-0.00	0.00**	0.01	-0.01	0.01	-0.00	-0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Com. language	-0.05	-0.02	-0.01	0.14*	-0.10***	-0.01	0.02	-0.00	0.01	0.13*	0.00	0.03	-0.08	0.01	-0.02	0.01	0.01	0.01
	(0.04)	(0.02)	(0.05)	(0.07)	(0.02)	(0.05)	(0.02)	(0.01)	(0.01)	(0.07)	(0.01)	(0.02)	(0.08)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
B-DOM	0.29	0.33**	0.23	0.66**	0.09	0.07	0.29***	-0.17	0.31***	0.67***	-0.06	0.13	0.13	0.22	0.17*	-0.03	-0.09	-0.00
	(0.25)	(0.16)	(0.21)	(0.30)	(0.48)	(0.19)	(0.07)	(0.11)	(0.09)	(0.25)	(0.14)	(0.17)	(0.35)	(0.14)	(0.09)	(0.17)	(0.09)	(0.08)
B-LESSVUL	-0.51*	0.05	-0.52**	0.14	0.48	0.09	0.03	-0.02	0.04	0.28	-0.03	0.24*	-0.35	0.11	-0.23*	-0.28*	0.12*	-0.05
	(0.27)	(0.07)	(0.26)	(0.73)	(0.48)	(0.43)	(0.08)	(0.07)	(0.06)	(0.23)	(0.07)	(0.15)	(0.31)	(0.10)	(0.14)	(0.16)	(0.07)	(0.09)
B-MOREVUL	-0.23	-0.11	-0.23	-0.49*	0.18	-0.43	0.08	-0.12	-0.02	0.25	-0.16**	0.13	-0.31	-0.07	-0.25**	-0.22	-0.05	-0.13
	(0.19)	(0.10)	(0.18)	(0.28)	(0.49)	(0.27)	(0.08)	(0.09)	(0.07)	(0.22)	(0.08)	(0.15)	(0.34)	(0.11)	(0.12)	(0.19)	(0.08)	(0.09)
B-NEA	-0.41**	-0.15	-0.32**	-0.75***	0.46***	0.08	-0.25***	0.21**	-0.25***	-0.36***	0.08	0.17	-0.47**	-0.06	-0.29***	-0.20*	0.23***	-0.03
	(0.17)	(0.15)	(0.12)	(0.27)	(0.07)	(0.09)	(0.05)	(0.08)	(0.08)	(0.12)	(0.12)	(0.11)	(0.20)	(0.12)	(0.05)	(0.11)	(0.08)	(0.04)
GG-DOM	0.01	0.59**	0.00	1.26**	-0.13	-0.05	-0.02	-0.09	-0.04	0.14	0.18	-0.03	0.03	0.21**	0.05	-0.02	-0.12	-0.30***
	(0.25)	(0.25)	(0.18)	(0.50)	(0.22)	(0.40)	(0.11)	(0.10)	(0.12)	(0.22)	(0.13)	(0.18)	(0.23)	(0.10)	(0.13)	(0.18)	(0.10)	(0.09)
GG-LESSVUL	-0.06	0.45*	-0.08	0.06	0.17	0.07	0.11	0.09	0.07	0.03	0.21**	-0.08	-0.23	0.24**	-0.01	0.15	0.15*	0.01
	(0.14)	(0.26)	(0.13)	(0.55)	(0.22)	(0.42)	(0.11)	(0.08)	(0.09)	(0.26)	(0.10)	(0.17)	(0.27)	(0.12)	(0.15)	(0.16)	(0.09)	(0.07)
GG-MOREVUL	-0.17	0.18	-0.20	-0.60	-0.14	-0.31	0.08	-0.07	-0.03	-0.22	-0.00	-0.36**	-0.28	0.02	-0.24	-0.07	-0.04	-0.27***
	(0.13)	(0.27)	(0.13)	(0.56)	(0.26)	(0.43)	(0.12)	(0.14)	(0.10)	(0.21)	(0.12)	(0.18)	(0.23)	(0.15)	(0.17)	(0.13)	(0.11)	(0.09)
GG-NEA	-0.02	-0.11***	0.01	-1.33***	0.39***	0.20	0.16***	0.21***	0.15*	-0.08	0.02	0.06	-0.06	0.07	-0.04	0.03	0.28***	0.23***
	(0.22)	(0.04)	(0.16)	(0.27)	(0.04)	(0.14)	(0.05)	(0.08)	(0.08)	(0.11)	(0.11)	(0.10)	(0.13)	(0.05)	(0.04)	(0.13)	(0.07)	(0.07)
NFC-DOM	-0.20	0.14	-0.20	2.44*	-0.20	0.39	-0.22*	-0.11	-0.12	0.40*	-0.09	0.31**	-0.67	0.19	-0.05	-0.03	-0.09	0.09
	(0.25)	(0.10)	(0.20)	(1.31)	(0.17)	(0.51)	(0.11)	(0.10)	(0.10)	(0.21)	(0.11)	(0.15)	(0.91)	(0.21)	(0.24)	(0.23)	(0.11)	(0.26)
NFC-LESSVUL	0.17	0.13*	0.13	0.18	0.13	0.20	-0.01	0.06	0.04	0.39**	-0.15***	0.30**	-0.10	0.19	-0.04	0.10	0.15	0.37
	(0.19)	(0.07)	(0.18)	(0.27)	(0.18)	(0.25)	(0.09)	(0.07)	(0.06)	(0.17)	(0.05)	(0.14)	(0.17)	(0.25)	(0.22)	(0.25)	(0.13)	(0.27)
NFC-MOREVUL	-0.14	0.05	-0.17	-0.37	-0.01	-0.50	0.08	0.07	0.08	0.41**	-0.12	0.33**	0.05	0.22	0.13	0.11	0.00	0.25
	(0.18)	(0.08)	(0.16)	(0.71)	(0.17)	(0.68)	(0.12)	(0.10)	(0.07)	(0.17)	(0.10)	(0.14)	(0.17)	(0.24)	(0.16)	(0.26)	(0.17)	(0.30)
NFC-NEA	0.09	0.02	0.09	-2.25*	0.35***	-0.17	0.17**	0.19***	0.14*	-0.08	-0.04	0.03	0.75	0.05	0.16	0.08	0.28***	0.23***
	(0.19)	(0.08)	(0.12)	(1.29)	(0.05)	(0.46)	(0.07)	(0.07)	(0.07)	(0.09)	(0.11)	(0.09)	(0.89)	(0.06)	(0.20)	(0.05)	(0.04)	(0.03)
Constant	0.94***	0.01	0.75***	0.90**	0.02	0.38*	0.16	0.05	0.28**	0.28	0.07	0.25*	0.66*	0.09	0.36***	0.31	-0.06	0.11
	(0.31)	(0.15)	(0.24)	(0.41)	(0.11)	(0.23)	(0.12)	(0.08)	(0.12)	(0.36)	(0.11)	(0.14)	(0.38)	(0.11)	(0.13)	(0.20)	(0.08)	(0.08)
Observations	1,712	1,329	2,153	1,073	1,902	2,150	2,408	3,462	3,824	1,277	2,458	2,631	806	1,918	2,077	1,162	3,180	3,309
R-squared	0.32	0.74	0.24	0.33	0.83	0.39	0.22	0.82	0.59	0.32	0.77	0.48	0.37	0.80	0.57	0.41	0.81	0.76

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

Covid Economics 65, 20 January 2021: 153-187

Table 5: Less vulnerable euro area countries - debt securities - 2020Q1-2020Q2

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h_{hc,hs}^{ic,a,is})$	-0.06**	-0.08***	-0.05***	-0.11**	-0.12***	-0.17***	-0.04***	-0.06***	-0.06***	-0.01	-0.02	-0.00	-0.10***	-0.05***	-0.05**	-0.01	-0.06***	-0.03***
	(0.02)	(0.02)	(0.02)	(0.04)	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Distance	-0.06	-0.03	-0.04	-0.07	-0.02	-0.10**	-0.02	-0.03**	-0.06***	0.07	0.01	-0.00	-0.08	0.00	-0.04	-0.02	-0.02	-0.01
	(0.05)	(0.02)	(0.03)	(0.09)	(0.02)	(0.04)	(0.02)	(0.01)	(0.02)	(0.07)	(0.02)	(0.02)	(0.08)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Imports	-0.02	0.01*	0.00	-0.02	0.01**	0.00	-0.02	0.01***	0.00	0.02	0.01*	0.01	-0.00	0.00	0.00	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.03)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Com. language	-0.07	0.05**	-0.08	0.14	-0.00	0.12	-0.02	0.00	-0.02	0.07	-0.00	-0.00	-0.01	0.01	-0.01	0.05	0.01	-0.00
	(0.06)	(0.02)	(0.06)	(0.20)	(0.02)	(0.08)	(0.04)	(0.01)	(0.02)	(0.09)	(0.02)	(0.02)	(0.13)	(0.02)	(0.04)	(0.05)	(0.01)	(0.01)
B-DOM	1.03***	0.13	0.86***	-1.43	0.08	0.04	0.85	0.07	0.24*	0.51	-0.05	-0.16	0.30	0.07	0.40	-0.08	-0.02	0.03
	(0.33)	(0.15)	(0.25)	(1.27)	(0.14)	(0.70)	(0.72)	(0.10)	(0.14)	(0.31)	(0.10)	(0.10)	(0.38)	(0.12)	(0.35)	(0.22)	(0.08)	(0.11)
B-LESSVUL	0.26	0.06	0.24	-1.08	0.45***	0.27	0.23***	0.07	0.11	0.10	-0.19	-0.29*	-0.22	0.02	0.04	-0.16	0.25***	0.16*
	(0.20)	(0.15)	(0.19)	(0.70)	(0.10)	(0.39)	(0.08)	(0.11)	(0.11)	(0.24)	(0.15)	(0.15)	(0.27)	(0.08)	(0.17)	(0.22)	(0.08)	(0.09)
B-MOREVUL	0.05	-0.01	0.10	-0.86*	0.23***	-0.43	0.05	0.03	-0.01	-0.19	-0.23	-0.29**	-0.46*	-0.01	-0.06	-0.01	0.18***	0.11*
	(0.31)	(0.08)	(0.21)	(0.45)	(0.05)	(0.62)	(0.08)	(0.11)	(0.09)	(0.21)	(0.14)	(0.15)	(0.23)	(0.07)	(0.09)	(0.06)	(0.06)	(0.06)
B-NEA	-0.44	0.08	-0.13	-0.04	0.50***	0.49***	-0.77	0.09	-0.13	-0.23	-0.04	-0.10	-0.46**	-0.01	-0.22	-0.01	0.27***	0.16**
	(0.29)	(0.12)	(0.10)	(0.47)	(0.10)	(0.17)	(0.72)	(0.12)	(0.15)	(0.19)	(0.14)	(0.14)	(0.21)	(0.10)	(0.27)	(0.10)	(0.08)	(0.08)
GG-DOM	-0.13	0.22	0.04	-1.80**	0.27**	0.53	0.03	0.03	0.08	0.14	-0.05	-0.14	-0.06	0.07	0.00	-0.42**	0.07	-0.05
	(0.54)	(0.14)	(0.35)	(0.84)	(0.11)	(0.43)	(0.15)	(0.11)	(0.13)	(0.33)	(0.13)	(0.11)	(0.66)	(0.09)	(0.18)	(0.17)	(0.13)	(0.13)
GG-LESSVUL	0.03	0.09	0.02	-0.48	0.50***	0.68***	0.26***	0.14	0.14	0.16	-0.08	-0.19	-0.57	0.09	-0.04	-0.07	0.33***	0.19***
	(0.30)	(0.12)	(0.28)	(0.62)	(0.07)	(0.24)	(0.07)	(0.10)	(0.11)	(0.26)	(0.14)	(0.14)	(0.39)	(0.07)	(0.13)	(0.10)	(0.08)	(0.07)
GG-MOREVUL	0.10	0.02	0.12	-1.12	0.36***	0.38**	0.16**	0.04	0.11	-0.01	-0.20	-0.25	-0.33*	-0.02	-0.14	-0.04	0.15	0.06
	(0.18)	(0.14)	(0.15)	(0.68)	(0.08)	(0.18)	(0.07)	(0.12)	(0.11)	(0.16)	(0.15)	(0.16)	(0.18)	(0.08)	(0.14)	(0.10)	(0.09)	(0.09)
GG-NEA	0.06	-0.09**	0.03	1.20**	0.35***	0.35**	0.11	0.16	0.09	0.15	0.02	-0.01	-0.20	0.03	-0.01	0.25***	0.23**	0.17
	(0.48)	(0.04)	(0.19)	(0.47)	(0.07)	(0.15)	(0.09)	(0.11)	(0.11)	(0.17)	(0.14)	(0.13)	(0.34)	(0.04)	(0.07)	(0.08)	(0.10)	(0.11)
NFC-DOM	-0.54	0.21	-0.30	0.41	0.31**	1.05	-0.11	0.09	0.09	0.41	0.09	0.02	-0.77	0.20**	0.12	-0.20*	0.11*	-0.02
	(0.48)	(0.16)	(0.34)	(0.48)	(0.14)	(0.83)	(0.14)	(0.08)	(0.10)	(0.34)	(0.10)	(0.09)	(1.42)	(0.08)	(0.43)	(0.11)	(0.07)	(0.06)
NFC-LESSVUL	-0.16	0.22*	-0.10	-0.56	0.58***	0.51***	0.37***	0.25**	0.26**	0.29	0.02	-0.07	-0.24	0.20***	0.23	-0.05	0.41***	0.27***
	(0.27)	(0.13)	(0.20)	(0.49)	(0.08)	(0.20)	(0.07)	(0.10)	(0.11)	(0.26)	(0.14)	(0.14)	(0.28)	(0.07)	(0.18)	(0.09)	(0.07)	(0.07)
NFC-MOREVUL	0.17	0.05	0.25	-0.87*	0.37***	0.32***	0.47	0.26**	0.25*	0.30	0.12	0.09	-0.35	0.12**	0.07	0.01	0.37***	0.28***
	(0.38)	(0.08)	(0.33)	(0.47)	(0.06)	(0.11)	(0.31)	(0.12)	(0.15)	(0.39)	(0.18)	(0.18)	(0.28)	(0.06)	(0.08)	(0.12)	(0.07)	(0.07)
NFC-NEA	0.18	0.08	0.13	-0.98**	0.38***	-0.33	0.32***	0.18*	0.15	0.01	-0.05	-0.10	0.89	0.03	0.32	0.13**	0.27***	0.21***
	(0.45)	(0.10)	(0.24)	(0.46)	(0.07)	(0.70)	(0.05)	(0.10)	(0.09)	(0.10)	(0.13)	(0.13)	(1.37)	(0.04)	(0.32)	(0.05)	(0.05)	(0.05)
Constant	0.84	0.20	0.46	1.48	0.11	0.93**	0.30	0.28*	0.67***	-0.63	0.08	0.17	1.22	0.10	0.50**	0.17	0.15	0.16
	(0.56)	(0.16)	(0.31)	(1.04)	(0.17)	(0.44)	(0.25)	(0.14)	(0.20)	(0.67)	(0.17)	(0.18)	(0.83)	(0.12)	(0.22)	(0.25)	(0.11)	(0.11)
Observations	714	758	971	446	1,152	1,177	959	1,644	1,660	524	1,143	1,149	470	1,216	1,235	540	1,511	1,525
R-squared	0.51	0.85	0.41	0.48	0.89	0.53	0.35	0.78	0.65	0.61	0.74	0.73	0.48	0.86	0.63	0.60	0.87	0.86

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

Covid Economics 65, 20 January 2021: 153-187

Table 6: More vulnerable euro area countries - debt securities - 2020Q1-2020Q2

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.09***	-0.05*	-0.06***	-0.16	-0.05*	-0.13	-0.04**	-0.03	-0.03**	0.01	-0.03	-0.05***	-0.13***	-0.05***	-0.06**	-0.03	-0.04**	-0.02
	(0.03)	(0.03)	(0.02)	(0.12)	(0.02)	(0.09)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance	-0.11	0.04	-0.06	-0.02	-0.00	-0.21	-0.05	0.02	-0.02	0.09	0.01	0.02	-0.30**	0.02	-0.06*	-0.10	0.04*	0.01
	(0.07)	(0.05)	(0.06)	(0.22)	(0.03)	(0.17)	(0.05)	(0.02)	(0.03)	(0.06)	(0.01)	(0.03)	(0.14)	(0.03)	(0.03)	(0.08)	(0.02)	(0.02)
Imports	-0.02	0.02	-0.01	0.02	0.00	-0.00	-0.00	0.01	0.01	-0.03**	0.01	0.00	0.04*	0.02	-0.00	0.04	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.07)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)
Com. language	0.06	0.02	0.04	0.01	-0.04	-0.05	-0.04	-0.01	-0.04	-0.15	0.02**	0.11	-0.39	0.00	0.00	-0.07	0.01	-0.00
	(0.15)	(0.03)	(0.11)	(0.20)	(0.06)	(0.13)	(0.11)	(0.01)	(0.04)	(0.12)	(0.01)	(0.07)	(0.31)	(0.02)	(0.03)	(0.10)	(0.02)	(0.02)
B-DOM	0.48	0.72*	0.57	0.19	0.20	-0.02	0.34**	0.07	0.40***	0.21	0.10	0.22	0.22	0.51	0.41***	-0.28	0.52**	0.24
	(0.31)	(0.44)	(0.39)	(1.11)	(0.13)	(0.43)	(0.17)	(0.13)	(0.11)	(0.22)	(0.12)	(0.19)	(0.32)	(0.33)	(0.13)	(0.52)	(0.21)	(0.15)
B-LESSVUL	0.19	-0.07	0.24	0.12	-0.14*	-0.38	-0.18**	0.14	0.18	-0.02	-0.12***	-0.14**	-0.52*	-0.13***	-0.22***	-0.22	0.28*	0.13
	(0.27)	(0.07)	(0.26)	(0.89)	(0.07)	(0.36)	(0.08)	(0.16)	(0.12)	(0.11)	(0.04)	(0.06)	(0.27)	(0.05)	(0.07)	(0.14)	(0.15)	(0.15)
B-MOREVUL	-0.20	-0.06	-0.05	-0.09	-0.11	-0.71	-0.13	0.12	0.25**	0.16	-0.12	-0.15	-0.87	-0.20**	-0.31*	-0.72	0.29***	0.17**
	(0.24)	(0.13)	(0.20)	(0.80)	(0.12)	(0.56)	(0.14)	(0.10)	(0.11)	(0.15)	(0.12)	(0.10)	(0.53)	(0.09)	(0.16)	(0.50)	(0.10)	(0.09)
B-NEA	-0.48**	-0.32	-0.47	0.36	-0.17**	-0.52	-0.51**	0.30***	-0.03	-0.19*	-0.00	-0.18	-0.72	-0.32	-0.58***	-0.21	0.13	0.11*
	(0.23)	(0.39)	(0.35)	(0.86)	(0.09)	(0.36)	(0.20)	(0.11)	(0.09)	(0.10)	(0.12)	(0.13)	(0.56)	(0.29)	(0.21)	(0.33)	(0.18)	(0.07)
GG-DOM	0.51	0.40	0.46	0.73	-0.12	-0.24	-0.03	-0.04	0.09	-0.02	-0.09	0.38	-0.20	0.15	0.00	0.07	0.16	0.04
	(0.44)	(0.31)	(0.33)	(1.41)	(0.18)	(0.82)	(0.17)	(0.19)	(0.18)	(0.28)	(0.16)	(0.24)	(0.54)	(0.22)	(0.19)	(0.37)	(0.18)	(0.16)
GG-LESSVUL	-0.47**	-0.07	-0.32*	0.75	-0.06	0.06	-0.20**	0.22	0.19	0.28	-0.04	0.11	-1.05***	-0.12**	-0.25***	-0.30*	0.34***	0.20
	(0.22)	(0.10)	(0.18)	(0.92)	(0.10)	(0.42)	(0.09)	(0.13)	(0.13)	(0.19)	(0.06)	(0.15)	(0.32)	(0.06)	(0.09)	(0.16)	(0.13)	(0.12)
GG-MOREVUL	0.16	-0.13	0.21	-0.74	-0.18	-1.15	-0.08	0.11	0.29**	0.15	-0.21**	0.01	-0.71	-0.22***	-0.37***	-0.24	0.23*	0.13
	(0.20)	(0.12)	(0.17)	(1.06)	(0.12)	(0.80)	(0.09)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)	(0.43)	(0.08)	(0.14)	(0.18)	(0.14)	(0.13)
GG-NEA	-0.40	-0.07	-0.26	-0.10	0.09	-0.40	-0.01	0.38***	0.36***	0.10	0.10	-0.16	0.04	-0.00	-0.15	-0.02	0.41***	0.32***
	(0.35)	(0.07)	(0.26)	(0.54)	(0.13)	(0.56)	(0.06)	(0.06)	(0.04)	(0.09)	(0.16)	(0.14)	(0.19)	(0.11)	(0.18)	(0.19)	(0.07)	(0.05)
NFC-DOM	0.32	0.34	0.08	1.92	0.07	1.05	-0.34*	0.11	0.04	0.02	0.11	0.31	-0.71	0.25	-0.13	-0.03	0.48**	0.22
	(0.38)	(0.30)	(0.27)	(1.71)	(0.17)	(1.23)	(0.18)	(0.15)	(0.12)	(0.25)	(0.13)	(0.20)	(0.49)	(0.22)	(0.13)	(0.47)	(0.21)	(0.16)
NFC-LESSVUL	-0.12	0.07	0.02	-0.18	0.04	-0.41	-0.01	0.30**	0.36***	0.21**	0.05	0.04	-0.46	0.00	-0.09	-0.13	0.43***	0.30**
	(0.26)	(0.09)	(0.22)	(1.12)	(0.08)	(0.44)	(0.10)	(0.14)	(0.13)	(0.11)	(0.06)	(0.06)	(0.28)	(0.05)	(0.08)	(0.14)	(0.14)	(0.13)
NFC-MOREVUL	-0.06	0.11	-0.07	1.04	0.13	0.03	-0.13	0.37***	0.30***	0.38***	0.13	0.11	-0.53	0.05	-0.11	-0.18	0.49***	0.37***
	(0.40)	(0.13)	(0.25)	(1.60)	(0.14)	(0.89)	(0.12)	(0.11)	(0.10)	(0.14)	(0.13)	(0.09)	(0.33)	(0.07)	(0.11)	(0.27)	(0.12)	(0.11)
NFC-NEA	-0.41*	0.01	-0.09	0.37	-0.01	-0.32	0.14***	0.27***	0.28***	0.08	0.02	-0.20*	0.27*	-0.08	-0.10	0.03	0.33***	0.26***
	(0.24)	(0.09)	(0.16)	(0.90)	(0.10)	(0.40)	(0.03)	(0.07)	(0.04)	(0.06)	(0.13)	(0.11)	(0.15)	(0.08)	(0.12)	(0.14)	(0.07)	(0.07)
Constant	1.63**	-0.43	0.91	0.39	0.14	2.48	0.82*	-0.32	0.08	-0.48	0.04	0.23	2.87**	0.08	0.96**	0.72	-0.56**	-0.21
	(0.80)	(0.49)	(0.65)	(1.99)	(0.31)	(1.74)	(0.48)	(0.21)	(0.24)	(0.54)	(0.22)	(0.31)	(1.30)	(0.30)	(0.43)	(0.90)	(0.26)	(0.21)
Observations	416	427	547	190	343	378	523	798	844	386	621	661	193	582	602	355	970	988
R-squared	0.54	0.70	0.60	0.48	0.92	0.48	0.44	0.95	0.68	0.47	0.93	0.60	0.69	0.90	0.79	0.58	0.88	0.85

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

6.3 Listed shares: shock period

6.3.1 All euro area countries

Starting with all euro area countries, Table 7 shows a significant "reversion to the mean" effect in listed shares (as also found for debt securities). However, in the case of listed shares, it is only significant for shifts in total exposures (for banks, OFI and PF). Moreover, distance matters less than for debt as it only affected the total holdings of IC (via direct exposures).

In terms of issuing sectors, only banks and NFC are relevant for listed shares. For both issuing sectors, there is evidence of a rebalancing away from domestic towards extra-euro area securities, which is the opposite of the general trend observed for debt securities. Banks are the only holding sector for which this rebalancing is exclusively driven by shifts in direct holdings. For other sectors reducing exposures to domestic banks and domestic NFC is at least partly driven by shifts in indirect holdings (IC, PF, NFC, HH). At the same time, there was a strong rebalancing by all sectors (except for banks) towards extra-euro area bank shares, sometimes solely driven by indirect holdings (OFI, NFC and HH). Even more so, this is visible in the rebalancing towards non-euro area shares issued by NFC, to which exposures increased for all sectors, which was driven by indirect holdings for banks, OFI, NFC and HH.

The rebalancing towards non-euro area listed shares may have been driven by the better relative stock market performance in non-European indices (in particular, in the US) and the fact that, in the second half of the first quarter of 2020, Europe was at the centre of the pandemic's developments with strong containment measures being enacted. What is more, this behaviour is in line with Broner et al. (2006), who find that mutual funds, when facing returns below average, tend to retrench from stocks of countries in which they are positioned overweight.

6.3.2 Less vulnerable euro area countries

Table 8 presents the equivalent results for investors from the less vulnerable euro area countries. Interestingly, for these countries there was a broad-based rebalancing by all sectors (except for banks) into the listed shares of domestic banks (i.e. the opposite of what was found in Table 7 for the entire universe of euro area investors) and into shares issued by banks from other less vulnerable euro area countries. In both cases these are almost exclusively driven by indirect holdings. At the same time, all sectors also increased their exposures to shares of banks resident outside the euro area, which for OFI, IC, NFC and HH was entirely driven by indirect holdings. Banks, on the other hand, increased their exposure to banks' shares issued by more vulnerable euro area countries, while OFI and IC reduced their exposure to these

Covid Economics 65, 20 January 2021: 153-187

CEPR PRESS

institutions.

Investors from the less vulnerable euro area countries, strongly rebalanced into listed shares issued by non-domestic NFC. For banks, OFI, IC, NFC and HH this was exclusively driven by indirect holdings, while for PF also direct holdings contributed. At the same time, HH and IC rebalanced away from domestic NFCs (if indirect holdings are included).

6.3.3 More vulnerable euro area countries

Table 9 shows the rebalancing in listed shares during 2020Q1 for the more vulnerable euro area countries. As regards bank shares, IC, PF and HH rebalanced away from domestic banks (due to indirect holdings), while the same sectors plus NFC increased their exposures to less vulnerable euro area and extra-euro area bank shares.

All these sectors also rebalanced into non-domestic NFC listed shares (both euro area and extra-euro area). As for the less vulnerable countries, this was largely due to indirect holdings, in particular for IC and NFC.

6.4 Listed shares: rebound period

In listed shares, the overarching patterns observed in the first quarter are also visible in the rebound period. The results for all euro area countries (Table 10) show that the large majority of sectors continued to increase their exposures to non-euro area listed shares (both issued by banks and NFC) at the expense of domestic listed shares. These patterns are once more driven to a large degree by indirect holdings, in particular for NFC and HH. At the country group level, the results are very much in line, with only few changes compared to the first quarter.[11]

[11] These results are available upon request.

Covid Economics 65, 20 January 2021: 153-187

Covid Economics 65, 20 January 2021: 153-187

Table 7: All euro area countries - listed shares - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.15***	-0.09***	-0.11***	-0.08***	-0.04*	-0.06***	-0.07***	-0.01	-0.01*	-0.03*	-0.01	-0.01	-0.07***	0.00	0.00	-0.03**	0.00	0.00
	(0.05)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance	-0.03	-0.09***	-0.01	-0.10	-0.01	-0.02	-0.07**	-0.00	-0.01	-0.02	0.01	-0.01	-0.13**	0.00	0.01	-0.04	-0.00	-0.01
	(0.16)	(0.04)	(0.07)	(0.08)	(0.01)	(0.05)	(0.04)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.05)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Imports	0.00	0.00	-0.00	-0.02	-0.01**	0.00	-0.01	-0.00	-0.00	0.00	0.01**	0.00	0.00	0.01*	-0.00	0.01	-0.00	0.01*
	(0.04)	(0.01)	(0.02)	(0.03)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Com. language	0.03	-0.18**	0.09	-0.00	0.14***	0.14**	0.10*	0.01	0.00	0.00	0.00	0.01	-0.02	-0.04**	0.06	0.02	0.00	-0.00
	(0.34)	(0.07)	(0.17)	(0.10)	(0.03)	(0.05)	(0.05)	(0.01)	(0.01)	(0.05)	(0.01)	(0.02)	(0.09)	(0.02)	(0.04)	(0.03)	(0.01)	(0.01)
B-DOM	-2.34***	-0.30*	-0.80**	-0.60	0.29**	-0.15	-0.09	-0.00	-0.04	-0.47***	-0.01	-0.34***	-0.19	-0.25***	-0.23	0.94***	-0.31***	-0.19***
	(0.84)	(0.16)	(0.40)	(0.66)	(0.12)	(0.41)	(0.19)	(0.05)	(0.07)	(0.15)	(0.06)	(0.13)	(0.27)	(0.05)	(0.14)	(0.10)	(0.05)	(0.06)
B-LESSVUL	-2.46**	-0.30**	-1.11*	-0.92	-0.19	-0.94**	0.29	0.20***	0.10	0.08	0.13*	-0.10	-0.68**	-0.10	-0.16	-0.24**	-0.06	-0.15*
	(1.07)	(0.15)	(0.60)	(0.57)	(0.12)	(0.38)	(0.27)	(0.06)	(0.08)	(0.18)	(0.06)	(0.13)	(0.29)	(0.07)	(0.15)	(0.11)	(0.06)	(0.08)
B-MOREVUL	-2.20**	-0.14	-0.68*	-0.48	-0.03	-0.49	0.22	0.29***	0.12	-0.13	0.18***	-0.18	-0.23	-0.03	0.02	-0.21**	0.02	-0.09**
	(0.90)	(0.12)	(0.39)	(0.55)	(0.15)	(0.42)	(0.21)	(0.06)	(0.08)	(0.17)	(0.04)	(0.13)	(0.33)	(0.05)	(0.15)	(0.09)	(0.03)	(0.05)
B-NEA	0.76***	0.15	0.43**	-0.37*	-0.39***	-0.48***	0.40***	0.24***	0.24***	0.51***	0.26***	0.26***	-0.08	0.25***	0.26***	-0.89***	0.26***	0.24***
	(0.10)	(0.11)	(0.19)	(0.21)	(0.11)	(0.12)	(0.06)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)	(0.13)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)
NFC-DOM	-0.52	-1.80***	-1.22***	0.15	0.01	0.89**	-0.53***	-1.33***	-1.28***	-0.61***	-1.18***	-1.24***	0.84*	-0.37***	-0.38**	0.10	-1.38***	-1.34***
	(0.95)	(0.36)	(0.42)	(0.52)	(0.13)	(0.39)	(0.18)	(0.10)	(0.12)	(0.21)	(0.04)	(0.08)	(0.47)	(0.08)	(0.16)	(0.10)	(0.07)	(0.09)
NFC-LESSVUL	-0.52	0.14	0.05	0.19	-0.24*	0.39	-0.14	0.05	0.07	-0.03	-0.03	-0.05	0.37	-0.04	-0.05	-0.13	-0.05	-0.11*
	(0.88)	(0.19)	(0.28)	(0.33)	(0.13)	(0.37)	(0.13)	(0.06)	(0.07)	(0.13)	(0.03)	(0.06)	(0.38)	(0.05)	(0.18)	(0.10)	(0.05)	(0.06)
NFC-MOREVUL	-0.52	0.29	0.14	0.23	-0.09	0.52	-0.17	0.07	0.04	-0.16	-0.03	-0.13*	0.36	-0.04	0.16	-0.25	0.01	-0.07
	(0.85)	(0.23)	(0.27)	(0.34)	(0.14)	(0.39)	(0.19)	(0.06)	(0.07)	(0.12)	(0.07)	(0.07)	(0.26)	(0.09)	(0.16)	(0.20)	(0.07)	(0.07)
NFC-NEA	0.75	1.93***	1.68***	-0.06	-0.17***	-0.21***	0.40***	1.39***	1.38***	0.65***	1.22***	1.20***	-0.44	0.38***	0.39***	-0.04	1.34***	1.34***
	(0.47)	(0.23)	(0.32)	(0.34)	(0.05)	(0.08)	(0.10)	(0.09)	(0.09)	(0.17)	(0.03)	(0.04)	(0.44)	(0.04)	(0.06)	(0.06)	(0.06)	(0.05)
Constant	-1.23	-0.03	-0.81	0.85	0.10	0.30	-0.26	-0.71***	-0.68***	-0.74**	-0.84***	-0.65***	0.75	-0.80***	-0.77***	0.50*	-0.79***	-0.70***
	(1.38)	(0.29)	(0.64)	(0.85)	(0.22)	(0.48)	(0.33)	(0.06)	(0.08)	(0.36)	(0.07)	(0.13)	(0.58)	(0.12)	(0.21)	(0.27)	(0.05)	(0.09)
Observations	353	616	723	515	900	1,010	587	1,787	1,821	496	1,238	1,306	423	955	1,022	838	1,669	1,798
R-squared	0.54	0.88	0.35	0.41	0.94	0.63	0.49	0.92	0.88	0.66	0.95	0.90	0.58	0.88	0.60	0.44	0.94	0.84

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by NFCs.

Covid Economics 65, 20 January 2021: 153-187

Table 8: Less vulnerable euro area countries - listed shares - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h_{hc,hs}^{ic,a,is})$	-0.18**	0.01	-0.11***	-0.06	0.04***	-0.09***	-0.09***	-0.04**	-0.04***	-0.03	-0.04***	-0.02	-0.07**	-0.05***	-0.01	-0.05*	-0.03	-0.02*
	(0.08)	(0.04)	(0.04)	(0.04)	(0.01)	(0.03)	(0.04)	(0.02)	(0.01)	(0.04)	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)
Distance	-0.04	-0.02	0.06	-0.03	0.01	0.05	-0.12*	-0.01	-0.02	0.12	-0.00	0.00	-0.20**	-0.02	-0.02	-0.05	-0.02*	-0.01
	(0.30)	(0.03)	(0.12)	(0.16)	(0.01)	(0.05)	(0.06)	(0.01)	(0.02)	(0.20)	(0.01)	(0.02)	(0.08)	(0.02)	(0.02)	(0.06)	(0.01)	(0.01)
Imports	-0.06	0.01	0.02	-0.00	0.00	0.01	-0.04	0.00	0.00	0.01	0.01	0.01*	-0.01	0.01	0.01	-0.00	-0.00	0.00
	(0.10)	(0.01)	(0.03)	(0.03)	(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.02)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Com. language	0.33	-0.13***	-0.08	-0.06	-0.00	0.11	0.10	-0.00	-0.06	0.03	0.02	0.03	0.04	-0.07***	-0.02	0.00	-0.02	-0.01
	(0.51)	(0.05)	(0.32)	(0.19)	(0.02)	(0.09)	(0.10)	(0.02)	(0.04)	(0.11)	(0.03)	(0.03)	(0.11)	(0.02)	(0.03)	(0.04)	(0.02)	(0.02)
B-DOM	-0.25	0.33	1.32*	0.71	0.21**	1.10***	-1.23**	0.32***	0.29**	0.62	0.35***	0.27**	0.12	0.37***	0.40***	0.87***	0.21***	0.54***
	(1.91)	(0.26)	(0.68)	(0.78)	(0.09)	(0.32)	(0.60)	(0.09)	(0.11)	(0.91)	(0.12)	(0.10)	(0.36)	(0.08)	(0.09)	(0.23)	(0.06)	(0.13)
B-LESSVUL	0.40	0.40***	0.72	0.46	0.22**	-0.10	-1.46**	0.45***	0.35***	0.91	0.40***	0.41***	-0.25	0.48***	0.40***	-0.25	0.45***	0.54***
	(1.54)	(0.15)	(0.51)	(0.72)	(0.09)	(0.30)	(0.63)	(0.06)	(0.08)	(0.92)	(0.06)	(0.06)	(0.34)	(0.05)	(0.06)	(0.25)	(0.04)	(0.12)
B-MOREVUL	1.73**	0.23	0.93**	-1.23***	-0.08*	-1.35***	-1.25***	-0.16***	-0.14***	0.74	0.04	0.11**	-0.36	0.35***	0.24***	-0.84***	-0.06	0.01
	(0.74)	(0.14)	(0.37)	(0.46)	(0.05)	(0.22)	(0.27)	(0.05)	(0.05)	(0.49)	(0.07)	(0.05)	(0.25)	(0.06)	(0.09)	(0.11)	(0.05)	(0.04)
B-NEA	0.82***	0.16	0.49***	-0.27	0.02	-0.57***	-0.17	0.25***	0.25***	0.61***	0.29***	0.32***	-0.21	0.28***	0.28***	-0.89***	0.28***	0.24***
	(0.17)	(0.14)	(0.19)	(0.35)	(0.07)	(0.17)	(0.24)	(0.02)	(0.02)	(0.14)	(0.04)	(0.04)	(0.14)	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)
NFC-DOM	0.64	0.22	1.16**	0.30	0.23***	1.00***	1.15***	-0.32**	-0.34***	0.84	-0.24*	-0.43***	0.36	0.33***	0.24	0.13	-0.47***	-0.45***
	(1.56)	(0.21)	(0.54)	(0.93)	(0.09)	(0.28)	(0.28)	(0.14)	(0.12)	(1.05)	(0.14)	(0.13)	(0.78)	(0.08)	(0.16)	(0.15)	(0.16)	(0.10)
NFC-LESSVUL	1.30	0.53***	1.10**	0.23	0.25***	0.16	-0.29	0.73***	0.71***	1.35	0.75***	0.65***	-0.33	0.79***	0.51***	-0.15	0.70***	0.61***
	(0.98)	(0.18)	(0.47)	(0.60)	(0.05)	(0.16)	(0.22)	(0.08)	(0.07)	(0.82)	(0.07)	(0.08)	(0.36)	(0.08)	(0.10)	(0.10)	(0.08)	(0.05)
NFC-MOREVUL	1.16*	0.78***	1.04***	0.08	0.42***	-0.19	-0.57*	0.73***	0.71***	0.96**	0.73***	0.70***	-0.13	0.79***	0.67***	-0.97	0.67***	0.37
	(0.68)	(0.14)	(0.33)	(0.37)	(0.07)	(0.26)	(0.34)	(0.07)	(0.06)	(0.40)	(0.07)	(0.08)	(0.21)	(0.07)	(0.07)	(0.72)	(0.07)	(0.24)
NFC-NEA	0.90	0.36**	0.55*	-0.04	0.00	-0.26**	-1.65***	1.13***	1.13***	0.73**	1.14***	1.23***	-0.87	0.57***	0.45***	-0.16**	1.19***	1.21***
	(0.66)	(0.15)	(0.29)	(0.45)	(0.04)	(0.11)	(0.31)	(0.04)	(0.05)	(0.34)	(0.05)	(0.04)	(0.54)	(0.06)	(0.07)	(0.07)	(0.06)	(0.04)
Constant	-0.72	-0.59**	-1.62	0.05	-0.53***	-0.17	1.15*	-0.65***	-0.55***	-2.11	-0.78***	-0.87***	1.62*	-0.63***	-0.62***	0.67	-0.60***	-0.67***
	(2.78)	(0.28)	(1.07)	(1.52)	(0.15)	(0.57)	(0.64)	(0.13)	(0.17)	(1.90)	(0.15)	(0.15)	(0.84)	(0.14)	(0.21)	(0.53)	(0.10)	(0.12)
Observations	218	362	406	285	534	571	239	851	857	208	576	581	257	593	615	417	780	818
R-squared	0.60	0.87	0.48	0.56	0.98	0.77	0.62	0.88	0.86	0.86	0.96	0.96	0.66	0.88	0.84	0.46	0.94	0.87

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by NFCs.

CEPR PRESS

Table 9: More vulnerable euro area countries - listed shares - 2019Q4-2020Q1

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h_{hc,hs}^{ic,a,is})$	-0.24**	-0.09***	-0.06	-0.09	0.06***	-0.06**	-0.08	0.00	-0.00	0.02	0.03	0.00	-0.11*	0.01	-0.02	-0.06***	-0.02	0.01
	(0.11)	(0.02)	(0.04)	(0.09)	(0.02)	(0.03)	(0.06)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.06)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)
Distance	0.27	-0.07*	-0.01	-0.46	0.07	-0.08	-0.20	-0.01	-0.05**	-0.00	0.03*	-0.01	-0.22	0.01	-0.06	-0.16*	0.00	-0.02
	(0.78)	(0.04)	(0.17)	(0.55)	(0.06)	(0.11)	(0.26)	(0.02)	(0.02)	(0.11)	(0.01)	(0.03)	(0.25)	(0.04)	(0.06)	(0.09)	(0.01)	(0.02)
Imports	-0.04	-0.05***	0.04	-0.17	-0.05**	-0.02*	0.05	-0.02***	-0.03***	0.00	-0.01***	-0.01	0.03	-0.01*	-0.02	0.01	-0.01***	0.00
	(0.34)	(0.02)	(0.07)	(0.14)	(0.02)	(0.01)	(0.08)	(0.01)	(0.01)	(0.02)	(0.00)	(0.01)	(0.06)	(0.00)	(0.02)	(0.02)	(0.00)	(0.01)
Com. language	-0.40	0.14***	-0.39*	0.37	0.07	0.04	0.31	0.02	0.04	-0.12	0.02**	-0.01	0.04	0.02	-0.00	-0.06	0.04*	0.02
	(0.29)	(0.05)	(0.23)	(0.41)	(0.07)	(0.07)	(0.29)	(0.02)	(0.03)	(0.08)	(0.01)	(0.05)	(0.20)	(0.02)	(0.06)	(0.09)	(0.02)	(0.02)
B-DOM	1.04	-1.09***	-0.03	-2.54	-0.39	-0.20	-0.10	-0.47***	-0.53***	-0.14	-0.10	-0.30	0.00	-0.30**	-0.29	-0.39	-0.31***	-0.20
	(6.01)	(0.24)	(1.08)	(2.90)	(0.29)	(0.29)	(1.42)	(0.09)	(0.11)	(0.44)	(0.07)	(0.21)	(1.24)	(0.12)	(0.26)	(0.46)	(0.07)	(0.19)
B-LESSVUL	-1.00	-0.73***	0.04	-0.62	0.29**	-0.45	0.13	0.52***	0.55***	0.22	0.16**	-0.09	-0.54	0.05	-0.19	0.42	0.54***	0.40***
	(0.99)	(0.15)	(0.66)	(1.09)	(0.12)	(0.31)	(0.81)	(0.04)	(0.13)	(0.32)	(0.08)	(0.13)	(0.45)	(0.13)	(0.14)	(0.51)	(0.04)	(0.05)
B-MOREVUL	-0.79	-0.95***	-0.48	-0.44	-0.17	-0.54	-0.65	0.01	-0.06	-0.29	-0.11*	-0.52***	-0.79	-0.42**	-0.63***	-0.02	0.03	0.01
	(1.89)	(0.15)	(0.84)	(0.82)	(0.15)	(0.35)	(0.61)	(0.04)	(0.07)	(0.32)	(0.06)	(0.18)	(0.56)	(0.17)	(0.20)	(0.48)	(0.07)	(0.04)
B-NEA	-0.25	-0.32***	0.41	-0.26	-0.04	-0.57**	-0.09***	0.22***	0.18***	-0.08	-0.11	-0.22***	-0.25	-0.20	-0.42***	0.62	0.26***	0.23***
	(0.73)	(0.10)	(0.73)	(0.52)	(0.09)	(0.25)	(0.02)	(0.02)	(0.02)	(0.21)	(0.08)	(0.08)	(0.18)	(0.13)	(0.12)	(0.48)	(0.01)	(0.03)
NFC-DOM	1.61	-0.22	0.56	-2.99	0.00	-0.02	0.13	0.05	-0.14	0.89**	0.20*	0.20	0.38	0.15	-0.12	-0.24	-0.64***	-0.93***
	(5.81)	(0.25)	(1.05)	(2.92)	(0.32)	(0.32)	(1.03)	(0.10)	(0.13)	(0.44)	(0.11)	(0.18)	(1.06)	(0.18)	(0.30)	(0.32)	(0.11)	(0.30)
NFC-LESSVUL	0.98	0.11	0.67	0.14	0.30***	0.16	0.07	0.69***	0.67***	0.36*	0.24***	0.23***	0.09	0.18***	0.12	0.83*	0.80***	0.54***
	(1.37)	(0.08)	(0.71)	(0.94)	(0.10)	(0.39)	(0.36)	(0.08)	(0.11)	(0.20)	(0.02)	(0.05)	(0.40)	(0.06)	(0.10)	(0.47)	(0.08)	(0.11)
NFC-MOREVUL	0.52	-0.18	0.46	-0.47	0.30**	-0.20	0.03	0.65***	0.60***	0.31	0.23***	0.13	0.02	0.09	-0.00	0.62	0.72***	0.57***
	(1.76)	(0.12)	(0.78)	(1.02)	(0.12)	(0.29)	(0.64)	(0.06)	(0.09)	(0.32)	(0.09)	(0.12)	(0.57)	(0.16)	(0.14)	(0.47)	(0.06)	(0.08)
NFC-NEA	0.04	-0.44***	0.36	0.71**	-0.09	-0.22	0.24	0.38***	0.38***	-0.56**	-0.06*	-0.08*	0.16	-0.11*	-0.16**	0.98*	1.30***	1.41***
	(0.39)	(0.07)	(0.63)	(0.29)	(0.06)	(0.19)	(0.86)	(0.05)	(0.07)	(0.28)	(0.03)	(0.05)	(0.67)	(0.07)	(0.07)	(0.52)	(0.04)	(0.05)
Constant	-1.16	0.88**	-0.95	5.53	-0.30	1.11	1.19	-0.56***	-0.08	-0.61	-0.66***	-0.23	1.62	-0.40	0.61	0.42	-0.64***	-0.53**
	(9.39)	(0.44)	(1.99)	(6.37)	(0.58)	(1.13)	(2.81)	(0.17)	(0.25)	(1.05)	(0.18)	(0.30)	(2.72)	(0.41)	(0.67)	(1.00)	(0.09)	(0.26)
Observations	95	191	227	76	166	185	109	402	404	166	311	326	99	281	286	188	497	511
R-squared	0.84	0.93	0.72	0.89	0.96	0.87	0.72	0.98	0.96	0.76	0.99	0.88	0.73	0.95	0.72	0.71	0.99	0.96

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by NFCs.

Covid Economics 65, 20 January 2021: 153-187

Table 10: All euro area countries - listed shares - 2020Q1-2020Q2

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h_{hc,hs}^{ic,a,is})$	-0.10**	0.01	-0.06**	-0.07***	-0.09***	-0.04***	-0.04**	0.01	-0.00	-0.00	-0.01	0.00	-0.04***	-0.01	-0.02***	-0.03***	0.01	0.00
	(0.04)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Distance	-0.02	-0.00	-0.11	-0.05	-0.03	-0.03	-0.04	0.00	-0.01	-0.03	-0.00	-0.02	-0.05	0.00	-0.04**	-0.03*	0.00	-0.01
	(0.24)	(0.02)	(0.08)	(0.06)	(0.02)	(0.03)	(0.04)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Imports	0.04	0.00	0.00	-0.03	0.01	-0.01	-0.02*	-0.00	-0.00	0.01	-0.00	0.00	-0.02	0.00	-0.00	-0.01*	-0.00	-0.00
	(0.09)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Com. language	-0.37	-0.03	-0.23*	0.07	-0.17***	-0.06	0.09*	0.01	0.02**	0.24***	0.03**	0.11***	0.02	0.02	0.00	0.02	0.02***	0.02*
	(0.37)	(0.02)	(0.12)	(0.11)	(0.04)	(0.04)	(0.05)	(0.01)	(0.01)	(0.06)	(0.01)	(0.02)	(0.05)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
B-DOM	-0.76	0.07	-1.22***	0.73	0.62***	0.56	-0.27	-0.48***	-0.48***	-0.21*	0.49***	-0.36***	0.17	0.16***	-0.53***	0.04	0.32***	-0.43***
	(1.18)	(0.11)	(0.35)	(0.51)	(0.16)	(0.37)	(0.27)	(0.05)	(0.04)	(0.11)	(0.07)	(0.11)	(0.16)	(0.06)	(0.06)	(0.08)	(0.12)	(0.06)
B-LESSVUL	-0.90	0.78***	-0.42	-0.04	-0.05	0.46	0.78**	0.05	0.08*	0.05	1.02***	0.03	0.15	0.64***	0.03	0.06	0.78***	0.01
	(1.10)	(0.08)	(0.38)	(0.51)	(0.12)	(0.53)	(0.33)	(0.04)	(0.04)	(0.20)	(0.10)	(0.08)	(0.16)	(0.05)	(0.08)	(0.08)	(0.12)	(0.08)
B-MOREVUL	-0.85	0.62***	-0.59**	0.03	-0.24	0.02	0.68**	-0.09	-0.05	-0.14	0.87***	-0.01	0.08	0.58***	-0.13	0.02	0.69***	-0.04
	(0.98)	(0.11)	(0.29)	(0.43)	(0.20)	(0.34)	(0.30)	(0.12)	(0.10)	(0.20)	(0.14)	(0.09)	(0.17)	(0.14)	(0.12)	(0.07)	(0.09)	(0.09)
B-NEA	0.52***	0.68***	0.83***	-0.54***	-0.54***	-0.37***	0.99***	0.49***	0.51***	0.52***	0.52***	0.51***	-0.16**	0.58***	0.56***	0.00	0.46***	0.42***
	(0.19)	(0.06)	(0.12)	(0.21)	(0.09)	(0.09)	(0.06)	(0.02)	(0.03)	(0.02)	(0.10)	(0.09)	(0.06)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
NFC-DOM	-2.44	-1.66***	-3.10***	0.26	-0.48**	-0.78***	-1.47***	-1.18***	-1.26***	-0.62***	-0.45***	-0.49***	0.30	-0.73***	-0.87***	0.16**	-1.08***	-1.07***
	(1.92)	(0.19)	(0.73)	(0.60)	(0.19)	(0.18)	(0.18)	(0.18)	(0.18)	(0.13)	(0.10)	(0.11)	(0.23)	(0.11)	(0.09)	(0.07)	(0.13)	(0.13)
NFC-LESSVUL	-2.59	0.14	-1.14	-0.16	-0.05	-0.13	-0.26*	0.04	-0.05	-0.10	-0.02	-0.00	0.10	0.05	-0.13	-0.01	0.01	-0.00
	(1.77)	(0.09)	(0.71)	(0.32)	(0.18)	(0.17)	(0.13)	(0.04)	(0.05)	(0.08)	(0.03)	(0.07)	(0.11)	(0.10)	(0.14)	(0.06)	(0.05)	(0.05)
NFC-MOREVUL	-2.88	0.07	-1.43**	-0.38*	-0.25	-0.36***	-0.31**	0.04	-0.08*	-0.06	-0.03	-0.03	-0.03	0.03	-0.13	-0.09	-0.02	-0.08*
	(1.76)	(0.06)	(0.71)	(0.22)	(0.18)	(0.14)	(0.14)	(0.03)	(0.04)	(0.13)	(0.04)	(0.09)	(0.20)	(0.10)	(0.10)	(0.07)	(0.05)	(0.05)
NFC-NEA	0.43**	1.74***	1.93***	-0.47	0.49***	0.63***	1.17***	1.19***	1.18***	0.69***	0.43***	0.47***	-0.32	0.82***	0.85***	-0.22***	1.05***	1.01***
	(0.19)	(0.18)	(0.26)	(0.53)	(0.10)	(0.11)	(0.10)	(0.18)	(0.18)	(0.06)	(0.10)	(0.09)	(0.21)	(0.04)	(0.04)	(0.05)	(0.12)	(0.12)
Constant	-0.53	-0.60***	0.15	1.09*	0.82***	0.71**	-0.55*	-0.47***	-0.39***	-0.23	-0.49***	-0.37**	1.02***	-0.57***	-0.21	0.51***	-0.44***	-0.28***
	(2.45)	(0.18)	(0.75)	(0.57)	(0.22)	(0.30)	(0.32)	(0.07)	(0.07)	(0.32)	(0.11)	(0.16)	(0.39)	(0.11)	(0.16)	(0.17)	(0.06)	(0.08)
Observations	341	625	714	521	903	1,014	568	1,760	1,791	491	1,221	1,289	423	940	1,010	846	1,646	1,780
R-squared	0.50	0.88	0.36	0.33	0.82	0.45	0.31	0.83	0.80	0.47	0.90	0.73	0.42	0.88	0.71	0.53	0.89	0.81

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by NFCs.

7 Results: portfolio shifts during the 2015-2017 APP episode

The analysis presented in the previous section offered a granular and nuanced view of the rebalancing patterns observed amid the COVID-19 shock and in the subsequent "rebound" quarter, particularly highlighting investor heterogeneity and the large role played by indirect holdings via investment funds. The question, emerges if similar patterns can also be observed by "looking-through" the large-scale portfolio rebalancing observed in response to the ECB's Asset Purchase Programme (APP) in the period 2015Q1 to 2017Q4 (Coeure, 2017).

The ECB announced, on 18 March 2020, the Pandemic Emergency Purchase Programme (PEPP), to counter the risks to the monetary policy transmission mechanism and the deterioration of the outlook for the euro area posed by the COVID-19 shock outburst. The first purchases took place on 26 of March, very close to the end of the first quarter (four working days). Although PEPP is similar, in nature, to the Asset Purchase Programme (APP), announced in early-2015 to support the monetary policy transmission mechanism and provide the amount of policy accommodation needed to ensure price stability, there are, however, important differences. On the one hand, the universe of eligible assets under PEPP is wider, as it includes (i) securities issued by the Greek Government; (ii) non-financial commercial paper (iii) public sector securities with a residual maturity of ranging from 70 days up to a maximum of 30 years and 364 days, while the APP, initially, only included securities with 2-year residual maturity (later extended to one year, at end-2016). Further to these differences in the universe of eligible assets, PEPP also differs from APP in that, despite the fact that the benchmark allocation across jurisdictions is also the capital key in the ECB of the national central banks, purchases are, however, conducted in a more flexible manner. This allows for fluctuations in the distribution of purchase flows over time, across asset classes and among euro area jurisdictions.

Table 11 displays for all euro area countries the results for the change in holdings in the period from end-2014 to 2017Q4, i.e., immediately before the largest component of the APP – the Public Sector Purchase Programme (PSPP) – was announced and up until the tapering off of purchases, which started in January 2018. The results for the euro area as a whole underscore very clearly that euro area investors, almost across the board, increased their exposure to securities issued by non-residents, while in many instances they decreased their exposure to domestic securities, issued by the government sector and banks. The latter is broadly in line with studies which specifically focused on the portfolio balance effects of the APP – see, for instance, Bergant et al. (2020). Strikingly, the rebalancing towards foreign debt securities is heavily driven by indirect holdings for IC, PF, NFC and HH. Thus, the main vehicle to

Covid Economics 65, 20 January 2021: 153-187

increase exposures to extra-euro area debt during the APP period, was precisely the one used to reduce these exposures during the COVID-19 shock: indirect holdings via investment fund shares. These broad patterns are also visible among the more and less vulnerable country groups.[12]

8 Conclusion

This paper has provided an in-depth perspective of the portfolio shifts brought about by the severe financial strains unleashed by the COVID-19 shock. It showed the different ways in which euro area sectors were exposed to debt securities and listed shares at the end of 2019, and how they rebalanced their portfolios, in the first quarter of 2020, characterised by financial market turmoil, and, subsequently, in the second quarter, when market functioning was restored.

More generally, the analysis in this paper highlights the important role that investment funds play for the exposures of most euro area sectors, as can also be observed for the large-scale rebalancing during the APP period from 2015 to 2017. It stresses the importance of bypassing and looking through holdings of investment fund shares, in order to be able to proper identify the ultimate underlying exposures of investors, in particular in times of large financial shocks.

[12]These results are available upon request.

Covid Economics 65, 20 January 2021: 153-187

Table 11: All euro area countries - debt securities - 2014Q4-2017Q4

	(1) B_dir	(2) B_ind	(3) B_tot	(4) OFI_dir	(5) OFI_ind	(6) OFI_tot	(7) IC_dir	(8) IC_ind	(9) IC_tot	(10) PF_dir	(11) PF_ind	(12) PF_tot	(13) NFC_dir	(14) NFC_ind	(15) NFC_tot	(16) HH_dir	(17) HH_ind	(18) HH_tot
$ln(h^{ic,a,is}_{hc,hs})$	-0.31***	-0.35***	-0.31***	-0.41***	-0.30***	-0.40***	-0.35***	-0.36***	-0.36***	-0.30***	-0.32***	-0.26***	-0.41***	-0.41***	-0.44***	-0.20***	-0.25***	-0.29***
	(0.03)	(0.04)	(0.02)	(0.05)	(0.03)	(0.04)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)
Distance	-0.37***	-0.19***	-0.35***	-0.06	-0.16***	-0.27***	-0.23***	-0.12***	-0.21***	-0.32***	-0.05**	-0.16***	-0.23**	-0.21***	-0.28***	-0.23***	-0.09***	-0.16***
	(0.07)	(0.03)	(0.06)	(0.09)	(0.03)	(0.06)	(0.04)	(0.02)	(0.03)	(0.06)	(0.02)	(0.04)	(0.11)	(0.03)	(0.06)	(0.06)	(0.02)	(0.03)
Imports	-0.02	0.04***	-0.00	-0.07*	0.02***	-0.00	0.02	0.01	0.01	0.01	0.00	0.03*	0.04	0.01	0.01	-0.03	0.00	0.00
	(0.03)	(0.01)	(0.03)	(0.04)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.05)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Com. language	-0.05	-0.06	0.02	0.49***	-0.02	0.10	0.13	0.00	0.11**	0.08	-0.05	-0.07	-0.00	-0.10***	-0.12*	-0.00	-0.03	-0.07**
	(0.11)	(0.06)	(0.12)	(0.18)	(0.03)	(0.13)	(0.09)	(0.03)	(0.05)	(0.14)	(0.03)	(0.05)	(0.18)	(0.04)	(0.07)	(0.12)	(0.02)	(0.03)
B-DOM	0.52	1.34***	0.66	-2.59**	0.66***	-3.00***	0.38	-2.61***	-2.47***	-1.42***	-2.20***	-2.63***	2.31***	-1.49***	-2.85***	0.48	-0.34	-1.35***
	(0.83)	(0.38)	(0.82)	(1.07)	(0.24)	(0.65)	(0.75)	(0.49)	(0.53)	(0.38)	(0.46)	(0.62)	(0.81)	(0.25)	(1.09)	(0.53)	(0.97)	(0.19)
B-LESSVUL	-0.20	-0.08	-0.19	-1.01	-0.35	-1.40**	-0.97*	-1.03**	-0.41	-2.69***	-0.80***	-1.89***	-0.25	0.51	-1.42	0.72	0.96	0.06
	(0.79)	(0.38)	(0.81)	(1.06)	(0.23)	(0.64)	(0.57)	(0.46)	(0.46)	(0.28)	(0.29)	(0.38)	(0.72)	(0.32)	(1.08)	(0.52)	(0.97)	(0.28)
B-MOREVUL	-2.14**	0.27	-1.94**	-1.20	0.47**	-1.54**	-0.95*	-0.85	-0.23	-2.52***	-0.41	-1.43***	-0.83	0.92***	-1.67	0.13	1.68*	0.11
	(1.06)	(0.49)	(0.97)	(1.05)	(0.23)	(0.73)	(0.53)	(0.58)	(0.42)	(0.41)	(0.44)	(0.42)	(0.78)	(0.23)	(1.09)	(0.49)	(1.01)	(0.25)
B-NEA	-0.32	-0.86***	-0.27	2.07***	-0.74***	2.04***	-0.98*	1.95***	2.36***	-1.06***	1.80***	1.27**	-1.28***	2.21***	2.22***	0.17	1.50***	1.50***
	(0.40)	(0.23)	(0.30)	(0.46)	(0.15)	(0.20)	(0.57)	(0.21)	(0.40)	(0.19)	(0.39)	(0.51)	(0.43)	(0.13)	(0.26)	(0.26)	(0.14)	(0.12)
GG-DOM	0.34	2.49**	1.39***	-0.96	0.65***	-1.65***	1.19***	-1.86***	-1.91***	0.92***	-2.32***	-0.68	2.27***	0.89***	-0.56	-1.14**	0.02	-0.65*
	(0.73)	(1.22)	(0.44)	(0.66)	(0.20)	(0.51)	(0.43)	(0.52)	(0.50)	(0.35)	(0.51)	(0.52)	(0.84)	(0.23)	(0.59)	(0.48)	(0.23)	(0.34)
GG-LESSVUL	-0.28	1.02	-0.13	-0.74	-0.44**	-0.45	-1.01***	-0.26	-0.76*	-0.10	-1.10***	0.16	-0.83	2.70***	0.80	0.71	1.16***	0.84**
	(0.30)	(1.15)	(0.31)	(0.55)	(0.17)	(0.68)	(0.38)	(0.48)	(0.43)	(0.30)	(0.34)	(0.25)	(0.70)	(0.27)	(0.80)	(0.46)	(0.22)	(0.34)
GG-MOREVUL	-0.75*	0.41	-0.33	-1.39*	-0.30	-1.68**	-1.76***	-1.06	-1.13**	0.49	-1.36***	-0.04	-0.69	2.35***	0.11	-0.67	1.11***	-0.07
	(0.39)	(1.22)	(0.39)	(0.79)	(0.32)	(0.83)	(0.61)	(0.67)	(0.55)	(0.41)	(0.44)	(0.33)	(0.86)	(0.18)	(0.67)	(0.55)	(0.28)	(0.57)
GG-NEA	-0.40	-1.11***	-1.04***	1.28***	-0.99***	1.56***	-1.30***	1.79***	1.98***	-0.67***	1.52***	1.17**	-1.35***	1.88***	1.91***	1.39***	1.33***	1.38***
	(0.66)	(0.27)	(0.34)	(0.36)	(0.20)	(0.20)	(0.23)	(0.22)	(0.32)	(0.22)	(0.39)	(0.50)	(0.47)	(0.14)	(0.20)	(0.28)	(0.14)	(0.13)
NFC-DOM	1.93***	0.94***	0.66	-2.84*	0.93***	-1.68	3.23***	-1.69***	-0.46	1.43***	0.52	0.74	1.77*	0.52***	-0.53	0.23	0.29	0.46
	(0.72)	(0.34)	(0.84)	(1.47)	(0.19)	(1.25)	(0.92)	(0.25)	(0.57)	(0.49)	(0.57)	(0.67)	(0.99)	(0.19)	(0.86)	(0.55)	(1.13)	(0.62)
NFC-LESSVUL	-1.08	-1.22***	-0.72	-1.14	-0.69***	-0.10	-0.07	-0.94***	0.16	0.31	1.17***	0.92*	-0.92	1.88***	0.56	0.33	0.88	1.39**
	(0.72)	(0.35)	(0.77)	(1.50)	(0.23)	(1.25)	(0.37)	(0.25)	(0.49)	(0.51)	(0.45)	(0.48)	(0.84)	(0.28)	(0.90)	(0.50)	(1.14)	(0.63)
NFC-MOREVUL	-1.84***	-1.05***	-1.62**	-1.41	-0.35*	-0.40	-0.28	-0.87*	0.32	0.66	1.40***	1.21**	-1.50*	1.81***	0.55	-0.20	1.06	1.66***
	(0.67)	(0.33)	(0.74)	(1.53)	(0.18)	(1.30)	(0.36)	(0.46)	(0.50)	(0.50)	(0.45)	(0.54)	(0.85)	(0.13)	(0.85)	(0.49)	(1.14)	(0.63)
NFC-NEA	-2.91***	-1.48***	-1.13***	1.85***	-1.33***	1.61***	-2.72***	1.22***	1.19***	-1.03***	1.08***	0.57	-1.54***	1.63***	1.77***	-0.09	0.84***	0.95***
	(0.29)	(0.25)	(0.43)	(0.50)	(0.17)	(0.30)	(0.85)	(0.21)	(0.33)	(0.17)	(0.39)	(0.52)	(0.47)	(0.14)	(0.17)	(0.27)	(0.14)	(0.16)
Constant	4.51***	3.19***	4.47***	-1.07	3.42***	1.82***	4.09***	0.67**	1.29***	3.76***	0.08	1.13*	3.18**	0.76***	1.39**	2.36***	0.85***	1.50***
	(0.81)	(0.41)	(0.72)	(0.86)	(0.31)	(0.60)	(0.48)	(0.27)	(0.40)	(0.66)	(0.42)	(0.62)	(1.27)	(0.25)	(0.55)	(0.67)	(0.22)	(0.26)
Observations	1,353	1,155	1,776	769	1,349	1,571	1,612	2,504	2,794	954	1,824	2,006	643	1,532	1,648	981	2,500	2,624
R-squared	0.49	0.82	0.50	0.47	0.89	0.49	0.62	0.93	0.79	0.58	0.84	0.68	0.60	0.86	0.71	0.54	0.88	0.80

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses, clustered at holder and issuer country-sector level. B-DOM, B-LESSVUL, B-MOREVUL and B-NEA are dummy variables for exposures to debt securities issued by banks resident, respectively, in the domestic economy, other less vulnerable euro area countries, other more vulnerable euro area countries and non-euro area countries. GG-DOM, GG-LESSVUL, GG-MOREVUL, GG-NEA, NFC-DOM, NFC-LESSVUL, NFC-MOREVUL, NFC-NEA are equivalent dummy variables for securities issued by the general government sector and NFCs.

Covid Economics 65, 20 January 2021: 153-187

CEPR PRESS

A Data appendix

A.1 Country aggregates

- **Euro area less-vulnerable countries** – Austria, Belgium, Finland, France, Germany, the Netherlands
- **Euro area more-vulnerable countries** – Greece, Italy, Portugal, Spain

A.2 Issuer country reclassifications due to the lack of gravity variables

- **Switzerland** – includes Liechtenstein
- **United Kingdom** – includes Guernsey, Jersey, Isle of Man, Gibraltar and the Virgin Islands
- **United States** – includes Puerto Rico, Guam, American Samoa, the US Virgin Islands and Minor Outlying Islands

A.3 Sources of gravity variables

- **Bilateral distance and shared language** – Data are taken from the GeoDist database, compiled by the Centre d'Études Prospectives et d'Informations Internationales (CEPII) – see Mayer and Zignago (2011) for details
- **Bilateral imports** – Data are from Eurostat and are four-quarter moving sums.

Covid Economics 65, 20 January 2021: 153-187

References

Adrian, T. and H. S. Shin (2010). Liquidity and leverage. *Journal of financial intermediation 19*(3), 418–437.

Adrian, T. and H. S. Shin (2013). Procyclical leverage and value-at-risk. *The Review of Financial Studies 27*(2), 373–403.

Altavilla, C., M. Pagano, and S. Simonelli (2017). Bank exposures and sovereign stress transmission. *Review of Finance 21*(6), 2103–2139.

Bergant, K., M. Fidora, and M. Schmitz (2020). International Capital Flows at the Security Level – Evidence from the ECB's Asset Purchase Programme. IMF Working Papers 20/46, International Monetary Fund.

Bergant, K. and M. Schmitz (2020). Valuation Effects and International Capital Flows - Security-Level Evidence from Euro Area Investors. mimeo, European Central Bank and International Monetary Fund.

Boermans, M. A. and R. Vermeulen (2019). International investment positions revisited: Investor heterogeneity and individual security characteristics. *Review of International Economics 28*(2), 466–496.

Broner, F., A. Erce, A. Martin, and J. Ventura (2014). Sovereign debt markets in turbulent times: Creditor discrimination and crowding-out effects. *Journal of Monetary Economics 61*, 114–142.

Broner, F., A. Martin, and J. Ventura (2010). Sovereign risk and secondary markets. *American Economic Review 100*(4), 1523–55.

Broner, F. A., R. G. Gelos, and C. M. Reinhart (2006). When in peril, retrench: Testing the portfolio channel of contagion. *Journal of International Economics 69*(1), 203–230.

Carvalho, D. (2020). The portfolio holdings of euro area investors - looking through investment funds.

Coeure, B. (2017). The international dimension of the ecb's asset purchase programme. Speech by Benoit Coeure at the Foreign Exchange Contact Group meeting, 11 July 2017.

ECB (2015). Who holds what? new information on securities holdings. *Economic Bulletin 2*, 72–84.

Fidora, M., M. Fratzscher, and C. Thimann (2007). Home bias in global bond and equity markets: the role of real exchange rate volatility. *Journal of international Money and Finance 26*(4), 631–655.

Covid Economics 65, 20 January 2021: 153-187

Galstyan, V. and P. Lane (2013). Bilateral portfolio dynamics during the global financial crisis. *European Economic Review 57*(C), 63–74.

Galstyan, V., P. R. Lane, C. Mehigan, and R. Mercado (2016). The holders and issuers of international portfolio securities. *Journal of the Japanese and International Economies 42*, 100–108.

Galstyan, V. and A. Velic (2018). International investment patterns: the case of german sectors. *Open Economies Review 29*(3), 665–685.

Giofré, M. (2013). International diversification: Households versus institutional investors. *The North American Journal of Economics and Finance 26*, 145–176.

Lane, P. R. (2020a). The macroeconomic impact of the pandemic and the policy response. The ECB Blog, 4 August.

Lane, P. R. (2020b). The market stabilisation role of the pandemic emergency purchase programme. The ECB Blog, 22 June.

Lane, P. R. (2020c). The pandemic emergency: the three challenges for the ecb. Speech at the Jackson Hole Economic Policy Symposium, Federal Reserve Bank of Kansas City "Navigating the Decade Ahead: Implications for Monetary Policy", 27 August.

Mayer, T. and S. Zignago (2011). Notes on cepii's distances measures: The geodist database. Working Papers 2011-25, CEPII.

Mehl, A., M. Schmitz, and C. Tille (2019, March). Distance(s) and the Volatility of International Trade(s). CEPR Discussion Papers 13630, C.E.P.R. Discussion Papers.

Monti, P. and A. Felettigh (2008). How to interpret the cpis data on the distribution of foreign portfolio assets in the presence of sizeable cross-border positions in mutual funds-evidence for italy and the main euro-area countries. *Bank of Italy Occasional Paper* (16).

Roque, V. and M. C. Cortez (2014). The determinants of international equity investment: Do they differ between institutional and noninstitutional investors? *Journal of Banking & Finance 49*, 469–482.

Timmer, Y. (2018). Cyclical investment behavior across financial institutions. *Journal of Financial Economics 129*(2), 268–286.

Make it or break it: Vaccination intent at the time of Covid-19

Jacques Bughin,[1] Michele Cincera,[2] Kelly Peters,[3] Dorota Reykowska,[4] Marcin Zyszkiewicz and Rafal Ohmed[5]

Date submitted: 30 December 2020; Date accepted: 6 January 2021

This research updates early studies on the intention to be vaccinated against the Covid-19 virus among a representative sample of adults in 6 European countries (France, Germany, Italy, Spain, Sweden, and the UK) and differentiated by groups of "acceptors", "refusers", and " hesitant". The research relies on a set of traditional logistic and more complex classification techniques such as Neural Networks and Random Forest techniques to determine common predictors of vaccination preferences. The findings highlight that socio-demographics are not a reliable measure of vaccination propensity, after one controls for different risk perceptions, and illustrate the key role of institutional and peer trust for vaccination success. Policymakers must build vaccine promotion techniques differentiated according to "acceptors", "refusers", and "hesitant", while restoring much larger trust in their actions upfront since the pandemics if one wishes the vaccination coverage to close part of the gap to the level of herd immunity.

Covid Economics 65, 20 January 2021: 188-222

1 Université libre de Bruxelles, Solvay Brussels School of Economics and Management, iCite and ECARES, MachaonAdvisory, Fortino Capital, Antler and Portulans Institute.
2 Université libre de Bruxelles, Solvay Brussels School of Economics and Management, iCite and ECARES.
3 BEworks.
4 Neurohm.
5 Neurohm, and WSB University.

1. Introduction

The Covid-19 pandemic continues unabated across the world. Official infections in the US have reached 18 million by mid-December 2020, far below the herd immunity needed to end the pandemic of the coronavirus SARS-CoV-2.

Currently the strategy set up by many countries has been to push for non-pharmaceutical measures. This strategy, while effective at flattening the curve, risks a fast rebuild of the pandemic if measures are loosened. Many countries have witnessed this issue with the second wave in most of the European countries, and a third wave emerging in yet a few countries (e.g. South Korea).

The only serious alternative is vaccination. Since April, the race has been on, with hundreds of candidates. This unprecedented effort has already led to two major vaccines being authorized by public health authorities. The vaccine set up by Pfizer/BioNTech and Moderna does not use bits of the virus; rather it leverages messenger RNA to trigger the immune system to produce antibodies. The reported protective effectiveness looks to be very promising, above 90% (See e.g. Polack et al. 2020). This level is slightly lower than measles but much more effective than traditional vaccines like the flu, at just above 60% (e.g. Osterholm et al. 2012). The technique, further, may open a totally new era for human protection.

Nevertheless, vaccination coverage must be widespread if the hope is stopping the pandemic. This is far from a "done deal". For example, compliance with the anti-H1N1 vaccine in 2009, was notoriously low. In Europe where most countries followed the vaccination recommendations of the WHO, barely 15% of the population went to be vaccinated, except for the Netherlands which initiated an aggressive vaccination campaign (Blasi et al., 2012).

One major cause of reluctance is the newness of a vaccine. The Covid-19 case is no exception, adding on top, the "newness" of the technique and the fact that the virus went into production in less than one year versus often a decade in the past for other vaccines.

Covid Economics 65, 20 January 2021: 188-222

In Europe, side effects/safety of the vaccine have been reported to be the largest reason for hesitating to be vaccinated against the Covid-19 (Neumann-Bohme et al., 2020, or Karlsson et al., 2020).[1] But, aside from the uncertainty around the product itself, another key worrying trend is the rise of social movements against vaccination. The rise of activists against vaccination is well documented and not new (Kata, 2012) but is booming in the case of the Covid-19 through fake news and conspiracy theories relayed online via social media (Nguyen and Catalan, 2020).[2]

The individuals receptive to those conspiracy theories are usually people with a feeling of being lost or betrayed, which reinforces the importance of transparency and trust as a key channel of influence for vaccination success (Schwartz, 2020). In this research, we are abstracting from the pure « product » effect, already well covered in Neumann-Bohme et al. (2020), and Karlsson et al. (2020), and looks at factors such as health, job, or financial risks as well as multi-dimensional trust, e.g., trust towards other individuals, and institutions such as the government and the media.

As in Neumann-Bohme et al. (2020), our research is based on an online survey with 1,000 respondents per country, among various European countries, conducted in April 2020. Differences with their study are however important. First, both studies cover France, Germany, Italy, and the UK, but we include Spain and Sweden. Second, as we are relying on respondents' statements, we adjust survey answers, by response time. This leverages the neuroeconomics principle that response time is a strong indicator of attitude strength (see Fazio et al., 1989). As we correct for this response time, we essentially recalibrate responses based on strength of survey responses and avoid noises based on answers hesitancy (we find the information bias to be important in case of vaccination leading to a reallocation of 7 points of positive answers to the category of vaccine

[1] Vaccination risk can also be anchored to recent controversies too. For example, the Pandemrix vaccine caused big controversy due to its association with an increased risk of narcolepsy (Sarkanen et al., 2018). Another controversy happened around the vaccine against measles, mumps, and rubella, as it was wrongly suggested that the vaccine could drive autism.

[2] The consequence of vaccination decline is leading to the resurgence of virus, -otherwise instinct, with high morbidity risk, like measles. In the US for instance, the share of vaccinated kids for important diseases such as MMR, Varicella, or Polio, is 6 points of a percentage lower than a decade ago in Georgia and Arkansas, as reported by Statista, based on US health testing centres.

Covid Economics 65, 20 January 2021: 188-222

hesitancy). Third, we contrast *three* groups, those who would accept to be vaccinated, those who will refuse, and those who are hesitating; each of the groups against the two others, so that we can determine the commonality and impact of drivers in each group.

Fourth, when it comes to risk perception, the literature on vaccination intent has primarily looked at *health* risk. However, individuals also may trade-off social distancing over-vaccination as an additional choice, and may also consider other risks such as job and finance preservation as part of their decision to get vaccinated. Our survey is rather comprehensive in covering the extent of risk *perception* during the Covid-19 crisis and demonstrates that other risks than health are to be taken into account as determinants of vaccination intent.

We also look at various trust elements. Algan et al. (2017) have for instance demonstrated that third-party trust is an important factor of vote extremism and rejection of the public authorities. This element of people trust is rarely measured in assessing vaccination intent, but maybe a relevant driver to the extent that public authorities actively promote vaccination (Jovančević and Milićević, 2020).

Last, but not least, the decision-making process of people to get vaccination is likely not linear. Game theory models of vaccination such as in Choi and Shim (2020), demonstrate that vaccination Nash strategies are a negative step function of the recovery rate of infection, and below which vaccination propensity increases in function of the opportunity costs of not being vaccinated. The latter is for instance driven by fatality rate or lockdown duration. In such a case, we may need to use non-linear techniques to determine vectors of vaccination preferences. For robustness, we use and compare typical logit, but also more complex Random Forest and Gradient Boosting Machine classification machine learning techniques as predictors of vaccination preferences.

In a nutshell, our research provides five key findings:

Covid Economics 65, 20 January 2021: 188-222

1) The percentage of acceptance is 65%[3] or a level that *does not guarantee herd immunity against the Covid-19.*

2) The risk of being contaminated, and given infection, the risk of large morbidity risk drive preferences for vaccination, *in contrast to job/financial risk which increases vaccination hesitancy*.

3) Institutional trust (towards health care, government, and the media) has a major influence on vaccination outcomes. *Trust in the former has the greatest influence on the undecided; confidence in the media and the government rate play on crushing the rate of refusal.* In particular, the degree of exaggeration of the media and the versatility of government actions undermine the credibility of discourse and the understanding of the disease. Trust towards *peers* also matters for vaccination choice.

4) When we control for attitudes, *only age (over 64 years of age) and very low educational attainment emerge as robust socio-economic drivers of preferences*.

5) The citizen *who respects the rules of confinement is more likely to recommend her social circle to respect them as well as to be vaccinated*. In this sense, putting citizens on the right side of the debate is also critical.

The next section discusses the research background, and high-level statistics including the will to take the vaccine. Multi-variate results are then presented and discussed. The last section concludes.

2. Research background and statistics

2.1. Research objectives

This research is a part of an extensive multinational project aimed at understanding people's attitudes, emotions, and behaviors connected with the SARS-CoV-2 pandemic, and their consequences for the adoption of protective behavior. This follows from the established fact that risk perception intensity may support a behavioral change to limit virus exposure (Wise et al., 2020; Harper et al., 2020).

[3] This number is smaller than in Neumann-Bohme et al. (2020), as we especially reweighted the sample to account for time response.

Covid Economics 65, 20 January 2021: 188-222

Using the same data set of this paper, a companion paper (Bughin et al., 2020) concentrates on risk perception and non-pharmaceutical protection interventions (NPIs) including quarantine, social distancing, and extra hygiene adoption. Risks include health morbidity risks, but also, risks such as job and financial stability, psychological risk, and social risks, among others. Those risks are also clear drivers of NPI compliance. In this paper, we focus on *vaccination as another Covid-19 pandemic mitigation* strategy.

2.2. Data sampling and scope

Our focus is on six European countries: France, Germany, Italy, Spain, Sweden, and the UK, which are both representative of different socio-economic models (Esping-Andersen, 1999)[4], as well of different archetypes of policy responses to the Covid-19 crisis. Countries like France, Spain, and Italy have limited trust in their governments, exhibiting roughly half the trust expressed by Sweden, and those countries further got largely hit during the first wave of the pandemic, with citizens went centrally imposed to comply with very restrictive lockdown measures as a bold move to curb the pandemic.

The data collection was performed online[5], based on country representative samples for age (above 18 years old) and gender, and recruited via a panel agency in April 2020, with a total sample of more than 6,000 answers, or a minimum of 1,000 per country countries (see Table 1). Respondents got email invites and were informed about the study scope. The task of the respondents was to evaluate if they agree with the statements presented on the screen.[6] To avoid people being « forced » to respond, or respond with answers

[4] Esping-Andersen (1999) distinguishes a social-democratic regime to which the Nordic countries belong, a liberal regime that prevails mainly in Anglo-Saxon countries, and a conservative regime that applies to continental Europe (France, Belgium, Germany, Austria, and The Netherlands). A fourth type, the Latin regime, is a subvariant of the latter, which includes countries in Southern Europe (Italy, Greece, Spain, and Portugal). As part of the social-democratic regime, Swedish citizens are driven by a State focused on the egalitarian principle, but accountable for its actions. The model seems effective, as the Swedish population expresses the largest trust towards their governments within Europe.See Eurostat barometer 2015 for example; Trust - Our World in Data

[5] We would like to thank Neurohm and Syno for collecting the data in all the countries.

[6] See Appendix 1.

Covid Economics 65, 20 January 2021: 188-222

that are not reflective of actual behavior, each question was structured to respond, on a 3 point scale (yes, hesitant, no).

A caveat of surveys is the uncertainty of the fit between what people report and their actual attitudes and behaviors. This is critical in a study like this one, as results may lead to inadequate public policy implications. We thus apply response time measurement, and adjust data, in line with Fazio et al. (1989). The authors find a high correlation between report and actual behavior among people with fast reaction time when expressing their opinions. iCode Smart test was used to collect the data (Ohme et al., 2020), with response time (RT) collected for each answer. RT given with a latency lower than 500 milliseconds (ms) (suspected to be given randomly) or higher than 10,000 ms (suspected to have been given after distraction) were eliminated. In total, this amounts to only 0.96% of dubious responses.[7]

To account for individual differences in reaction speed, we standardize reaction time data measured in milliseconds, with STDRT being the z-score of log(RT), with mean = 0 and standard deviation = 1. For a question such as "do you envisage to get vaccinated', the proportion Y of citizens responding YES, is adjusted such that Y' = (1-a) x (Y) where (1-a) = max(SDRT, 2)/2. Thus $0 < a < 1$ acts as a factor that reduces the difference in positive responses in favor of hesitancy; in our sample, a = 19%, implying that gross responses on vaccination acceptance, at 72%, are over-rated by about 7 points.

[7] Furthermore, to ensure high quality of data and eliminate test biases a calibration phase and control screen have been added. Calibration preceded the test phase and consisted of 3 steps:

a. Familiarization with the scale. The task of the respondents was to press certain answer options – this task made sure respondents are aware of the position of the buttons on the screen.
b. Familiarization with the purpose of the task. A few statements were presented describing the test and the task. After each screen respondents had to press a button. This part served as a motoric warm up.
c. Increasing the focus on the task. During the study a screen appeared asking the respondent to indicate the statement that was presented last. The aim of this task was to make sure respondents focus their attention on the presented statements. Such screen was presented twice.

The control screen was introduced to eliminate the effect of the position of the mouse on the screen. It was presented before each statement, forcing a standardized position of the mouse (the distance to the yes and no answers was always the same).

Covid Economics 65, 20 January 2021: 188-222

2.3. High-level data statistics: Willingness to be vaccinated

Table 1 displays the number of respondents including the appetence to vaccination, after removing outliers.[8]

Table 1. Number of Covid-19 European respondents, April 2020

Countries	Total	Gender		Age	Willingness to be vaccinated		
	N	Females	Males	>64	Yes	Hesitant	No
France	1,024	51%	49%	13%	0.56	0.28	0.16
Germany	1,017	49%	51%	12%	0.63	0.20	0.17
Italy	1,021	51%	49%	10%	0.71	0.19	0.10
Spain	1,019	50%	50%	6%	0.74	0.18	0.08
Sweden	1,006	50%	50%	19%	0.56	0.28	0.16
The UK	1,068	53%	47%	24%	0.71	0.21	0.08
6 countries	6,155	51%	49%	14%	0.65	0.22	0.13

The total portion of citizens willing to be vaccinated is 65%, for about 22% "hesitant", and the balance (13%) are "refusers". The portion of "acceptors" is slightly less than other studies Neumann-Bohmer et al., 2020 estimate an intent at 74% for Europe, while in Lazarus et al. (2020), 80% would accept the Covid-19 vaccine. In our sample, we adjust for uncertainty in answers respondents, reducing positive direct responses on vaccination acceptance, at 72% by about 7 points.[9]

The difference in vaccination intent we observe per country is also visible in Neumann-Bohmer et al. (2020), with France, having the largest portion of refusers and hesitant.[10] In itself, the portion of acceptors is likely not to be sufficient to achieve herd immunity. This level is given by 1 - ($1/R_o$) where R_o is the basic reproduction number. If only to be estimated, R_o estimates range from 2.5 to 3.5, with a mean of 3.3 worldwide (Bughin,

[8] Outliers have been defined as responses with RT < 500 ms or RT > 10,000 ms. The first case implies that answers may only be random, the second case implies that people got distracted, with low confidence in the answer.

[9] People who hesitate to get vaccinated are probably more prone to see risks everywhere and trust less and respond slowlyer and more thoughtfully.

[10] Detoc et al. (2020) find a higher portion by aggregating those certainly and likely certainly to take the vaccine in France.

2020).[11] For Europe, the most recent estimates make it more towards 3.9 (Flaxman et al., 2020) and 3.4 for the US (Pizter et al., 2020). At those levels of estimates, herd immunity would need a level of acceptance of a minimum of 60%, and more likely up to 75%, *which is in practice challenging out of the figures of vaccination intent above.*[12]

2.4. High-level data statistics: Socio-demographics

The general consensus regarding vaccination reluctance is the mediating effect of some socio-economic factors such as the level of education, level of income, gender, and age (Larson et al., 2014). Neumann-Bohmer et al. (2020) find differences in the will to be vaccinated in gender, and in age, with a higher portion of males willing to get vaccinated, and of people about 55 years old. Table 2 provides a picture from our sample.

We also see that men have a higher propensity to accept vaccination, with lower hesitancy and refusal than women. Age exhibits a "U" curve regarding vaccination, with the lowest acceptance rate between 36-49 years old, and acceleration after 64 years old. Given retirement at an older age, retired citizens exhibit higher acceptance too. Lower education citizens have lower acceptance as is lower-income, following the literature. A-political citizens are much less inclined to be vaccinated, and stand for the largest hesitant group, with the unemployed. Interestingly, those citizens fit the category of the citizens feeling misfit with their socio-economic system, e.g. Algan et al. (2017), and belong to a growing category of believers of alternative, even, conspiracy theories.

In total, hesitancy ranges between 17% and 30%, refusal between 7% and 16% depending on the socio-economic cut. The best case (17% and 7%) leads to a maximum acceptance of 76%, which is barely the threshold of herd immunity at 90% rate effectiveness of the

[11] For comparison the 2002 SRAS' Ro was estimated to be in the range of 2.2 to 3.6. Ebola by 2014, is said to have a reproduction rate, Ro between 1.5 to 2.3. The 1918 Spanish influenza Ro was estimated imprecisely between 1.8 to 4.

[12] A way to get an idea of herd immunity is to look at extreme case studies. E.g., at San Quentin State Prison in California, more than 60% of the population was ultimately infected before the outbreak was halted. Lewis et al. (2020) report on the Brazilian city of Manaus which was devastated by a large outbreak of Covid-19. But by early June, the number of excess deaths from around 120 per day went to nearly zero. When researchers found that citizens with antibodies were reaching 66% of total population.

vaccine and for an $R_0 > 3$. This best-case however concentrates only on the retired, older age group, a group with the largest morbidity and mortality risks linked to the Covid-19, but also a small portion of the total European population.

Table 2. Socio-demographics of vaccination

When a COVID-19 vaccine is available I'd like to be vaccinated				
Typology	Details	Acceptor	Hesitant	Refuser
Gender	Female	0.62	0.25	0.13
	Male	0.69	0.19	0.12
Age	18-25	0.65	0.23	0.12
	26-35	0.61	0.24	0.15
	36-49	0.60	0.24	0.16
	50-64	0.66	0.23	0.11
	>64	0.78	0.15	0.07
education	Primary schools	0.62	0.25	0.13
	Middle school	0.64	0.22	0.14
	Vocational	0.62	0.23	0.15
	High school	0.67	0.22	0.11
	Bachelor or higher	0.67	0.22	0.11
kids	0 children	0.62	0.24	0.13
	1 child	0.68	0.19	0.13
	2 children	0.71	0.20	0.10
	3 children	0.63	0.21	0.16
	>3 children	0.60	0.25	0.15
occupations	Student	0.67	0.20	0.12
	Employed	0.64	0.23	0.13
	Entrepreneur	0.59	0.25	0.16
	Unemployed	0.60	0.27	0.14
	Retired	0.74	0.17	0.08
location	<100000 inhabitants	0.63	0.23	0.14
	>100000 inhabitants	0.65	0.21	0.13
income	<20000€	0.61	0.25	0.14
	>20000€	0.67	0.20	0.13
	don't want to answer	0.57	0.28	0.15
politics	Left	0.73	0.18	0.10
	Right	0.68	0.20	0.13
	Other	0.61	0.22	0.16
	Don't associate with politics	0.54	0.30	0.16
	Don't want to answer	0.62	0.26	0.11

2.5. High-level statistics: beyond socio-demographics

Appendix 2 further provides the univariate correlation matrix of respondents' statements from Appendix 1, with vaccination intent. The matrix gives a glimpse at the role of institutional trust and health risk, as vaccination intent is a) highly negatively correlated with factors such "Media exaggerate the situation with Covid-19" and b) highly positively correlated with "I am worried …about my health… about the health of

Covid Economics 65, 20 January 2021: 188-222

my children…about the health of my older family members”; as well as “Coronavirus is dangerous for my health”.

Heath and other Covid-19 risks

If agents are fully informed, voluntary vaccination arises when the relative benefit of vaccination becomes greater than the cost of vaccination. This relative benefit is greater when direct protection is greater (via high resistance to the pathogens) but is also lower when indirect protection is lower (via a lower portion of people being a transmitter of the disease). Thus, it implies that major health risks would correlate positively with the will to get vaccinated, as well as possibly with the perception of additional risk linked to lasting lockdown. In general, also, there is a typical “free rider” problem where vaccination propensity may marginally decline when more and more other individuals get the vaccine (Choi et al., 2020).

In our sample, we measure four types of worries mostly, health (henceforth, H), economic (E), social (S), and psychological (P). Table 3 provides the worries’ propensity ranked from the most frequently expressed worry to the lowest, and for the 16 constructs allocated to H, E, S, and P.

Except for the job situation, more than 1 out of 2 respondents worry about any matter. Health worries are the largest ones (average = 65.6%), while the lowest is financial worries (55.3%). Psychological and social risks are clearly important too (58% and 62% respectively).

We also describe the worries for the socioeconomic groups, with apparently lower vaccination propensity in Table 2. We notice that women are the most worried on all dimensions, and the reverse for the low-income segment. Lower education expresses lower worries than average except when it comes to their own and children’s health; while those in the 36-49 years old brackets generally are more worried than average, except for social violence. In general, those patterns do not demonstrate that worries correlate necessarily with lower vaccination expression by socio-demographic group, as it emerged from high-level data. Clearly other factors explain vaccination preferences.

Table 3. European citizens worry during the 1st wave of the covid-19 pandemic

		Vaccine "a priori less acceptors" groups				
Statement	**Yes**	**Women**	**Low education**	**36-49 years old**	**Low income**	**Do not associate with politics**
I am worried about the health of my older family members (H)	0.71	0.73	0.68	0.76	0.73	0.74
COVID-19 increases domestic violence (S)	0.64	0.66	0.66	0.63	0.64	0.62
The COVID-19 outbreak will make society more unequal (S)	0.61	0.61	0.60	0.63	0.63	0.62
I am worried that our country will run out of money (E)	0.63	0.66	0.58	0.65	0.66	0.67
I am worried about not being able to meet with my family (P)	0.63	0.65	0.62	0.65	0.64	0.68
COVID-19 will increase divorce rates (S)	0.60	0.61	0.60	0.61	0.60	0.61
I am anxious about not being able to meet with friends (P)	0.59	0.60	0.59	0.59	0.61	0.60
Living in isolation negatively impacts my wellbeing (P)	0.56	0.56	0.54	0.57	0.57	0.56
I am worried about my own health (H)	0.66	0.68	0.68	0.69	0.69	0.70
I am worried about the health of my children (H)	0.60	0.62	0.65	0.62	0.59	0.64
Being together all the time increases family tensions (S)	0.55	0.56	0.55	0.57	0.55	0.55
I worry how living in isolation will affect me (P)	0.55	0.57	0.53	0.58	0.58	0.58
I am worried about my financial situation (E)	0.56	0.59	0.52	0.62	0.64	0.61
I am worried about my job situation (E)	0.47	0.49	0.38	0.56	0.53	0.53

Note: all variables are adjusted for time response.

The role of institutional and peer trust

Under a new vaccine, like in this case of Covid-19, the assumption of perfect information on the effectiveness of the vaccine may evidently not hold. When uncertainty may then become a deterrent to accepting to be vaccinated, institutions may play an important role in transparency and information dissemination to support risk-averse people to take vaccination (Lazarus et al., 2020). This is even more important as there has been large documentation of cases of anti-vaccination groups promoting misinformation and

Covid Economics 65, 20 January 2021: 188-222

conspiracy theories, in order to create division, and aimed at sowing mistrust of experts' voices and government actions during a pandemic (Kata, 2010 and Fisher et al., 2020).

We have collected responses linked to trust around how the Covid-19 has been managed, along with three components, government/media, healthcare system, and people in general (Table 4).

Table 4. European citizens third party trust during the Covid-19 pandemic

Statement		Vaccine acceptor				
	Yes	Wo men	Low educa tion	36-49 years old	Low inco-me	Do not associate with politics
Trust in institutions						
I am satisfied with how my government is handling this crisis	0.59	0.59	0.60	0.57	0.58	0.51
The government is doing a good job dealing with Covid-19	0.58	0.59	0.61	0.57	0.57	0.51
The government discloses real numbers of coronavirus infections and deaths	0.54	0.53	0.54	0.53	0.53	0.49
[President] is doing a good job dealing with Covid-19	0.54	0.54	0.54	0.54	0.54	0.51
Media provide reliable information about the pandemic	0.57	0.58	0.57	0.57	0.57	0.54
Trust in Healthcare						
In case of coronavirus infection, I will get appropriate medical help	0.63	0.63	0.59	0.63	0.63	0.61
I am satisfied with how our healthcare system is handling this crisis	0.66	0.66	0.70	0.65	0.63	0.60
Trust in people						
Covid-19 reveals the best in people	0.57	0.57	0.58	0.56	0.56	0.55

Note: all variables are adjusted for time response.

Trust is far from being complete. This lack of trust likely is likely to drive down the acceptance of control measures, as already demonstrated in other pandemics (Gellin, 2020). Trust is larger towards healthcare than for institutions, and towards peers in general. Trust is notably low for the group that does not associate with politics, and except for the low-income segment, is rarely above the average for any other group, leading to the idea that low trust may build large hesitancy about vaccination.

Covid Economics 65, 20 January 2021: 188-222

3. Multi-variate analysis

3.1. Techniques used

We now resort to various formal multivariate analyses of potential acceptance of the Covid-19. The first technique is a simple logit regression model, the second is a classification tree model, the third one is a random forest classification model, and the last two ones are a Neuronal Network and a Gradient Boosting Machine classification models.[13] No model is better than the other, but we aim to have a comparison of multiple techniques for robustness.

In particular, it has been claimed that the Random Forest technique is important as an alternative to the traditional logit model, given likely non-linearity (e.g. threshold effect in health risk impact on control preferences), and high-order interaction effects. For example, Random Forest techniques have exhibited superior predictive power for H5N1 influenza outbreaks (see Herrick et al., 2013; or Kane et al., 2014), and recently Covid-19 infections (Iwendi et al., 2020).

This is still to be seen for vaccination choice, however. Table 5 presents a comparison of the fit accuracy of the various models, using a K-fold cross-validation (KFCV) technique (Li, 1987).[14] The accuracy is fair, at 79% for Random Forest, but at the same order of magnitude as other techniques.[15]

Covid Economics 65, 20 January 2021: 188-222

[13] The Random Forest algorithm is an ensemble learning method combined with multiple decision tree predictors that are trained based on random data samples and feature subsets (Breiman, 2001). Deep learning approaches are getting more and more used in complex non-linear predictive models. In the case of the Covid-19, it has been used successfully to predict Covid-19 infection development (Yeşilkanat, 2020), Covid-19 fatality rates (Pourhomayoun and Shakibi, 2020), or Covid-19 pulmonary damages (Sun et al., 2020).

[14] Formally, we keep aside a portion of the entire dataset, which is not used to train the models, and later use this sample for testing/validating the models. Formally, we split the dataset randomly into 10 folds, then fit the model using the K - 1 folds and validate the model using the remaining K^{th} fold. We repeat this process until every K-fold serves as the test set and then we take the average of the recorded scores for each validation.

[15] The total variance explained is 18.3%, and the mean square residuals is 18%.

Covid Economics 65, 20 January 2021: 188-222

Given likely non-linearity, at higher order, hierarchical effects, assumed, we focus here on Random Forest regression technique to discuss the results. We present a high-level comparison of the four techniques. Appendix 3a and 3b present the results of the logistic and tree regression for perusal examples.

Table 5. Performance (10-fold cross-validation) comparison of models

Model		Accuracy
Logistic regression on	acceptors	71.49%
	hesitant	77.38%
	refusers	87.08%
Classification tree on	acceptors	64.85%
	hesitant	76.59%
	refusers	86.85%
Random forest on	acceptors	70.56%
	hesitant	77.58%
	refusers	87.37%
Neuronal Network	acceptors	67.77%
	hesitant	72.51%
	refusers	83.88%
Gradient Boosting Machine	acceptors	71.54%
	hesitant	77.43%
	refusers	87.30%

3.2. Random forest tree results

We have configured the Random Forest algorithm with 5,000 trees in the forest, and including statements and demographics, 88 variables were tried at each split. Figures 1.1 to 1.3 exhibit the results for the three segments of "acceptors", "hesitant", and "refusers", ranked by decreasing importance on node purity (a measure of contribution fit) and on mean square error (a measure of predictive accuracy). The following insights emerge.

The first observation is that the two measures provide the same, but not fully equal, picture. As an illustration, Figure 2 ranks the top 20 most important factors (100% is the most important, and 0% the least important) for the acceptors. The rank fit exhibits a positive correlation, r = 0.83, with in general, the most important factors are slightly less predictive, and vice versa. Among factors relatively *more* predictive than anticipated by the pure regression fit, *those linked to attitude towards the healthcare system and healthcare professionals stand out, and rank among the top 10 most important factors to predict vaccination preferences*.

Figures 1a-1c. Random forest correlates and predictors of Covid-19 vaccination preferences (based on adjusted responses rate, RTC)

Figure 1.a. Random forest on acceptors

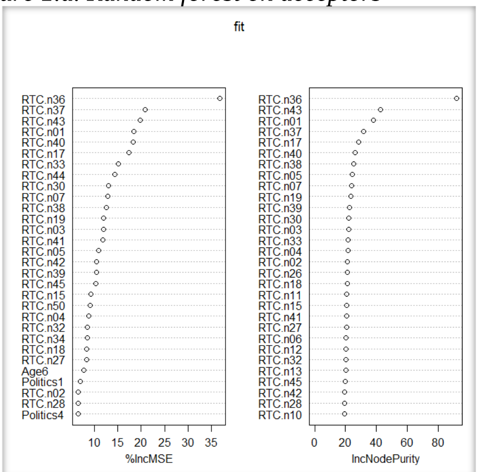


Figure 1.b. Random forest on hesitant

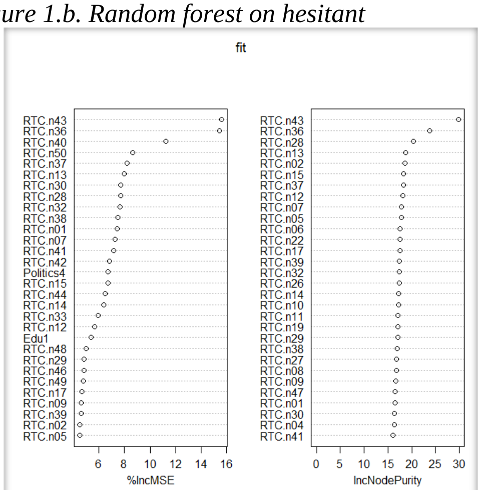
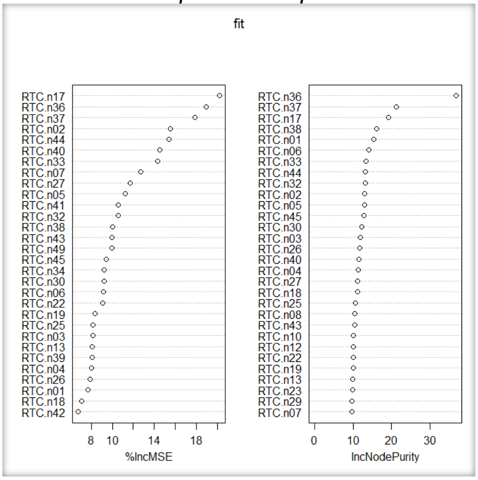
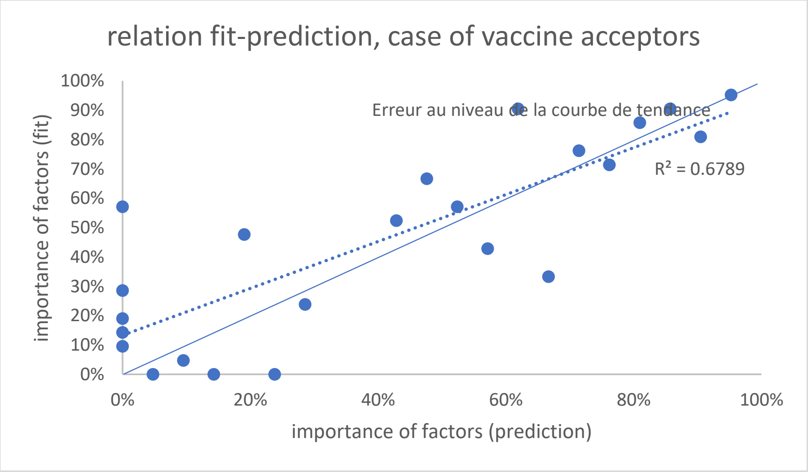


Figure 1.c. Random forest on refusers



The second observation is that the list of predictors is not the same for the different groups (will accept, hesitate to take, or refuse the vaccine). Ranking all factors by importance, the largest correlation observed is only r = 0.36 between acceptors and refusers. The correlation with the hesitant is neither significant for the acceptors, nor the refusers.

Figure 2. Comparison fit and prediction, Covid-19 vaccine acceptance



Covid Economics 65, 20 January 2021: 188-222

The hesitant are *more worried about their health* (and in accordance comply more strongly with social distancing measures), doubt more about media and government *transparency* than the two other groups, and on how covid may impact *their peers*. Regarding the refusers, they are much *less worried about their health* and their older family members' health than all other segments. They also are relatively *more worried about their personal and family tensions during Covid-19, and their finance*. And while they look at the government to be effective, they *consider the current actions to be largely not adequate*. Finally, regarding the acceptors, they are likely more prone to influence others to comply with control measures, they are confident in the healthcare treatment if infected while being worried about the virus.

Third, the importance of trust in predicting vaccination preferences is relatively material. We aggregate the top 20 most predictive factors along the dimensions of worries and trust, such that total = 100% in Table 6. *Trust factors in the aggregate are the largest predictors of vaccination preferences*. As we further split trust by constituents, and risk/worries by nature, health is evidently the most important predictor, but *institutional trust and psychological worries are also material*. As to be anticipated, health factors are relatively less important for refusers than for other vaccination preference categories, while refusers overweight economic worries and lower trust in people and government institutions. Hesitant overweight on psychological versus economic worries.

Fourth, only age (> 64 years) plays a predictive role in vaccination intent. This is indeed the age where the fatality risk linked to the Covid-19 increases significantly. Non-association to mainstream politics is also associated (but negatively) with vaccination acceptance, in line with this segment likely less trustful in institutions, as mentioned here-before. Low education predicts vaccination hesitancy, but no socio-economic variable seems to be exclusively linked to vaccine “refusers”.

Covid Economics 65, 20 January 2021: 188-222

Table 6. Comparison fit and prediction, Covid-19 vaccine preferences

Dimensions		Importance			
		total	Acceptors	Refusers	Hesitant
risk/worries	Health	46%	53%	29%	49%
	psychological	36%	27%	54%	46%
	economic	26%	17%	65%	0%
	total worries	37%	36%	40%	40%
trust	heathcare	80%	75%	49%	71%
	government	57%	57%	61%	55%
	Media	86%	86%	80%	81%
	people	27%	41%	63%	33%
	total trust	63%	65%	61%	60%

3.3.Comparison with other predictive models

As other models used (pure logistic, and regression trees) are as / if not more, informative as / than Random Forest, we complete this section by highlighting common results. Table 7 presents the top 10 common factors that come out as predictors of increased vaccination rates.

Table 7. Top 10 common factors that come out as determinants of increased vaccination rates

	Method (rank)			Effect on the will to vaccinate
statements	Random Forest	Regression Tree	Logistic	
Coronavirus is dangerous for my health	1	3	3	+
In the case of Covid-19, I will receive appropriate health	3	1	2	+
Media exaggerate the situation	2	4	1	-
I am satisfied with how the government handles the crisis	4	7	9	+
Government discloses real numbers of coronavirus	5	n.a	5	+
I am grateful to healthcare professionals	6	n.a	16	+
I actively ask people to follow measures against the virus	8	6	11	+
I am worried about my own health	11	2	4	+
I comply with measures of staying home	17	n.a	8	+
Covid-19 reveals the best in people	16	5	n.a.	+

Notes:
Rank based on predictive power averaged for “acceptors”, “refusers” and “hesitant”;
Only 7 statements are significant predictors in the case of a simple regression tree.

Covid Economics 65, 20 January 2021: 188-222

The common list of predictors is rather robust, with a rank correlation between 0.5 to 0.6 between pairs of statistical techniques. But differences also prevail. In practice:

1. *Health risk elements stand as a key opportunity cost* of not being vaccinated, as well as does health care quality- this again fits with existing literature and with the logic of optimal individual strategies for protective measures adoption (e.g. Choi and Shim, 2020)

2. *Trust also appears to be the key determinant of vaccination intent* in all techniques, as we anticipated. Among others, institutional trust is ranked higher within Random Forest than with the two other techniques. This is of major importance as the government effective prevention and rescue during the pandemic crisis are essential to get the full cooperation with their citizens (Zang, 2013, Li et al., 2018, Duan et al., 2020).

3. All techniques demonstrate *the importance of some individuals to actively seek others to comply with control measures.* This is consistent with other research highlighting that word of mouth is rather effective in previous vaccination campaigns (Tchuenche et al., 2011; Bhattacharyya et al., 2015).

4. Random Forest and logistic regression spot some extra factors, but not in the top 10 most important ones predicting vaccination preferences. Among others, Random Forest highlights the importance of the citizens' financial and personal situation. This shows the relevance of also securing wealth and social structure on top of health in a pandemic crisis.

5. Socio-demographics are not playing a major direct important role in predicting vaccination preferences, in any technique. This reinforces the point that attitudes and beliefs are much more crucial than anything others, to secure large vaccination potential.

Covid Economics 65, 20 January 2021: 188-222

3.4 . By the way, does all the above matter?

All the above points out that an important foundation that underpins vaccination acceptance is trust and sufficient consideration of the risks and worries of the population.

Here we use the estimates collected to see how big those effects are. If one takes for instance the results of the three regressions in Appendix 3, the best case leads to up to 12 extra points of vaccination intent, of which about 8 points is linked to trust improvement and tackling additional worries. Table 8 computes the average of the methods and shows that the total effect is about 14.5 points gross, and corrected for response time, 12 points extra.

At this indicative level, the vaccination rate may become close to 80% and may lead to close to the gap needed for herd immunity, if the effectiveness of vaccines is as expected to be above 90%, and vaccination intent fully materializes. This again re-emphasizes the importance of getting barriers removed for a future without Covid-19 morbidities.

Table 8. Marginal impact of fixing trust and other worries
Estimates on vaccination intent

Levers	marginal effect	population Coverage	total
Trust improvement	28%	31%	8.6%
Reduce media exageration	11%	39%	4.0%
Promote best in people	3%	29%	0.9%
Improve government actions	14%	26%	3.6%
fix additional worries	15%	41%	6.0%
Secure Job and finance	3%	50%	1.5%
secure appropriate medical help	12%	39%	4.5%
Total			14.5%

Source: Regression estimates weighted by nodes

3.5. Country effects

We close this section by reporting on country effects, in Table 9.

Covid Economics 65, 20 January 2021: 188-222

CEPR PRESS

Table 9. Country effects on vaccination preferences (after controlling for risk, trust, and sociodemographics)

Country	acceptors	hesitant	refusors
Spain	74%	19%	6%
France	57%	25%	17%
Italy	68%	26%	6%
Sweden	46%	30%	23%
Germany	77%	14%	9%
The UK			7%

Note: only significant effects at 5% are reported. Germany as the reference case.

Those effects in Table 9 suggest that Sweden is the least inclined to accept the vaccination but the will is much larger in countries most affected by the Covid-19 pandemic by April 2020, such as Spain and Italy. Remember that those are marginal probability estimates, i.e. after taking into consideration trust and perception risk. Sweden's perception risk has been actually lower than other countries, in the first wave of the Covid-19, given its limited lockdown strategy, even if Swedish citizens have had a large trust in its institutions. The reversal is true for Spain for example. In general, the figures suggest that country effects linked to vaccination preferences are rather strong and relatively larger than for non-pharmaceutical compliance propensity (Bughin et al., 2020).

4. Conclusions

This research has provided a view on European citizens' vaccination preferences for the SARS-CoV-2 and how those preferences are shaped by socio-demographics, risk perception of the pandemics, as well as trust in institutions and peers. Using multiple regression and classification techniques, all converge to the findings that vaccination preferences are shaped more by attitudes than by socio-demographics. In particular, it is critical to improving trust extensively in media, governments, healthcare, and peers if one hopes to be as close as to herd immunity with enough vaccination acceptance.

Covid Economics 65, 20 January 2021: 188-222

The research has some clear shortcomings, e.g. perceptions were collected in wave 1 of the pandemic, and may have changed since the evolution towards wave 2 (and even wave 3 in some countries), and the discovery and rollout of mRNA vaccines since late December 2020.

Furthermore, in our research, preferences were simply stated, but not forced against a plausible alternative like done in conjoint techniques. For example, asking whether people would want to be vaccinated or continue to be under lockdown, may lead to more optimistic results in favor of vaccination. We aim for those extensions in further research. Meanwhile, we believe our results have emphasized four core actions in the rollout of vaccination strategies.

1. **Build the urgency of vaccination.** The (adjusted) acceptance rate for vaccination (assuming the product is not controversial) is around 65%. Knowing that official contamination of the order of 5% of the population to date in Europe, the possible effective immunization rate would therefore be at best 64% under voluntary vaccination. This will lead to herd immunity if a basic Covid-19 reproduction rate, Ro < 2.8. While most estimates suggest Ro > 3 (and most likely Ro = 3.9 in Europe). Taking point estimates from other studies that uncertainty about the vaccine results could reduce the willingness to vaccinate, by 5-10 points (average 7), current voluntary vaccination may still imply (1 - (1/3.9)) - (64% - 7%) = 17% of the population still to be infected to get to herd immunity, or still three times the current level of infections.[16]
2. **Launch segment-specific campaigns on the no as well on the undecided.** Adjusting for response time in our survey, the non-vaccination rate is 12% while the portion of hesitation is 23%. In other words, the undecided are the most important class to convert. As seen above, the campaigns are however to be segmented. The hesitant must be reassured about how health institutions may support their own health and more transparency must be warranted by media

[16] Given asymptomatic cases, and absence of tracking, the figures of currnt infections may be two times higher. Than would still make the current shortfall to herd immunity to be as large as current infections, under no additional control measures.

Covid Economics 65, 20 January 2021: 188-222

and government alike. Refusers are relatively more worried about family tensions and finance, while they want to see governments that are much more effective in their actions, namely embracing more than health issues only.

3. **Reset the institutional factor**. Trust in Health care, government, and the media are more important than health risk perceptions if one wants to convert the undecided to accept the vaccine. In particular, it is crucial that media provides a true fact base on the pandemic challenges, while the government actions are simple, consistent in order to build the credibility of discourse
4. **Use the citizen force of persuasion**. In all cases, as in any social system, the citizen is also an important vector of influence. The importance of social media is often seen from the negative side, that is, social media may propagate fake news that increase the confusion in citizen mind, and thus reduce the will to get vaccinated. But the positive story is also that the citizen who is willing to vaccinate is also likely to recommend to her social circle to accept the vaccination. Given that acceptance is dominating today, the positive word of mouth must be much more systematically harnessed in order to make the undecided shift their mind, and possibly limit the contagion of refusers into the undecided population. For everyone's sake, and to get rid of the virus.

References

Algan, Y., S. Guriev, E. Papaioannou, E. Passari (2017). The European trust crisis and the rise of populism. Brookings Papers on Economic Activity, 2017(2), 309-400.

Bhattacharyya, S., C.T. Bauch, R. Breban (2015). Role of word-of-mouth for programs of voluntary vaccination: A game-theoretic approach. Mathematical biosciences, 269, 130-134.

Bhattacharyya, S., C.T. Bauch, R. Breban (2015). Role of word-of-mouth for programs of voluntary vaccination: A game-theoretic approach. *Mathematical biosciences*, *269*, 130-134.

Blasi, F., S. Aliberti, M. Mantero, S. Centanni, (2012). Compliance with anti-H1N1 vaccine among healthcare workers and general population. Clin. Microbiol. Infect. (18): 37–41.

Breiman, L. (2001). Random Forests. Machine Learning.

Bughin, J. (2020). Ten moments of truth in the Covid-19 Crisis Policy Punchline, Princeton.

Bughin, J., M. Cincera, R. Ohme, D. Reykowska, M. Żyszkiewicz (2020a). Perceptive risk clusters of European citizens and NPI compliance in face of the Covid-19 pandemics. Covid Economics 63: 126-158.

Choi, W., E. Shim (2020). Optimal strategies for vaccination and social distancing in a game-theoretic epidemiologic model. Journal of theoretical biology (505): 110422.

Detoc, M., S. Bruel, P. Frappe, B. Tardy, E. Botelho-Nevers, A. Gagneux-Brunon (2020). Intention to participate in a COVID-19 vaccine clinical trial and to get vaccinated against COVID-19 in France during the pandemic. Vaccine, 38(45), 7002-7006.

Duan, T., H. Jiang, X. Deng, Q. Zhang, F. Wang (2020). Government intervention, risk perception, and the adoption of protective action recommendations: Evidence from the COVID-19 prevention and control experience of China. International Journal of Environmental Research and Public Health, 17(10), 3387.

Esping-Andersen, G (1999). Social foundations of postindustrial economies, Oxford University Press.

Fazio, R. H (1989). The role of attitude accessibility in the attitude to-behavior process. The Journal of Consumer Research, 16(3), 280–288.

Fisher, K., B. Bloomstone, J. Walder, S. Crawford, H. Fouayzi (2020). Attitudes toward a potential SARS-CoV-2 vaccine: A survey of U.S. adults. Annals of Internal Medicine, M20-3569.

Covid Economics 65, 20 January 2021: 188-222

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T.A. Mellan, H. Coupland, M. Monod (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257-261.

Gellin, B. (2020). Why vaccine rumours stick—and getting them unstuck. *The Lancet*, *396*(10247), 303-304.

Harper C.A., L. Satchell, D. Fido, R. Latzman (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. Int. J. Ment. Health Addict (27): 1–14. 10.1017/dmp.2020.338

Herrick K.A., F. Huettmann, M. Lindgren (2013). A global model of avian influenza prediction in wild birds: the importance of northern regions. Veterinary Research. 2013, 44: 42-10.

Kane, M.J., N. Price, M. Scotch, P. Rabinowitz (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC bioinformatics, 15(1), 276.

Karlsson, L.C., A. Soveri, S. Lewandowsky, L. Karlsson, H. Karlsson, S. Nolvi, J. Antfolk (2020). Fearing the Disease or the Vaccine: The Case of COVID-19. Personality and Individual Differences, 110590.

Kata, A. (2012). Anti-vaccine activists, Web 2.0, and the postmodern paradigm - An overview of tactics and tropes used online by the anti-vaccination movement. Vaccine (30): 3778–3789.

Jovančević, A., N. Milićević (2020). Optimism-pessimism, conspiracy theories and general trust as factors contributing to COVID-19 related behavior–A cross-cultural study. *Personality and individual differences*, *167*, 110216.

Larson H, C. Jarrett, E. Eckersberger, D. Smith and P.Paterson (2014), Understanding vaccine hesitancy around vaccines from a global perspective : a systematic review of published literature, 2007-2012, Vaccine (32): 2150-59.

Lazarus, J.V., S.C. Ratzan, A. Palayew, L.O. Gostin, H.J. Larson, K. Rabin, A. El-Mohandes (2020). A global survey of potential acceptance of a COVID-19 vaccine. Nature medicine, 1-4.

Lewis, C., C. Prete, C. Abrahim, A. Mendrone Jr, T. Salomon, C. de Almeida-Neto, R. França, M. Belotti, M. Carvalho, A. Costa, M. Crispim, S. Fraiji, S. Gurzenda, Ch. Whittaker, L. Kamaura, P. Takecian, M. Oikawa, A. Nishiya ,V Rocha, N. Salles, A. de Souza-Santos, M. da Silva B Custe, M. Barral-Netto, M. Kraemer, R. Pereira, O. Pybus, M. Busch, M. Castro , Ch. Nascimento, N. Faria, E. Sabino (2020). COVID-19 herd immunity in the Brazilian Amazon, MedXriv, Sept 21.

Li, K.-C. (1987). Asymptotic optimality for Cp, CL, cross-validation and generalized cross-validation: discrete index set. Ann. Stat. 15(3): 958–975.

Li, Y.L.; W.Z. Wang, J. Wang (2018). Government intervention behavior and optimization strategy of major epidemic control: Based on game theory and China's H7N9 prevention and control practice in 2013. J. Hunan Agri. Uni. (Soc. Sci.), 19, 61–66.

Neumann-Bohme, S., N.E. Varghese, I. Sabat, P.P. Barros, W. Brouwer, J. van Exel, T. Stargardt (2020). Once we have it, will we use it? A European survey on willingness to be vaccinated against COVID-19. The European Journal of Health Economics (21): 977–982.

Nichols, T.R., P.M. Wisner, G. Cripe, L. Gulabchand (2011). Putting the Kappa Statistic to Use Thomas. Quality Assurance Journal, 57–61. https://doi.org/10.1002/qaj

Nguyen, A., D. Catalan (2020). Digital mis/disinformation and public engagement with health and science controversies: fresh perspectives from Covid-19. Media and Communication, 8(2), 323-328.

Ohme, R., M. Matukin, P. Wicher (2020). Merging Explicit Declarations With Implicit Response Time to Better Predict Behavior. In Chkoniya, V., Madsen, A. O., & Bukhrashvili, P. (Ed.), Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior (pp. 427-448). IGI Global.

Osterholm, M.T., N.S. Kelley, A. Sommer, E.A. Belongia (2012). Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. *The Lancet infectious diseases*, *12*(1), 36-44

Pitzer, V.E., M. Chitwood, J. Havumaki, N.A. Menzies, S. Perniciaro,J.L. Warren, T. Cohen (2020). The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *medRxiv*.

Polack, F.P., S.J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, R. Bailey (2020). Safety and efficacy of the BNT162b2 mRNA covid-19 vaccine. *New England Journal of Medicine*.

Pourhomayoun, M., M. Shakibi (2020). Predicting mortality risk in patients with COVID-19 using artificial intelligence to help medical decision-making. medRxiv.

Sarkanen, T., A. Alakuijala, I. Julkunen, M. Partinen (2018). Narcolepsy associated with Pandemrix vaccine. *Current neurology and neuroscience reports*, *18*(7), 43.

Schwartz, J. L. (2020). Evaluating and Deploying Covid-19 Vaccines—The Importance of Transparency, Scientific Integrity, and Public Trust. New England Journal of Medicine, 383(18), 1703-1705.

Sun, L., Z. Mo, F. Yan, L. Xia, F. Shan, Z. Ding, H. Yuan (2020). Adaptive feature selection guided deep forest for covid-19 classification with chest ct. IEEE Journal of Biomedical and Health Informatics, 24(10): 2798-2805.

Covid Economics 65, 20 January 2021: 188-222

Tchuenche, J. M., Dube, N., Bhunu, C. P., Smith, R. J., & Bauch, C. T. (2011). The impact of media coverage on the transmission dynamics of human influenza. *BMC Public Health*, *11*(S1), S5.

Wise T., T.D. Zbozinek, G. Michelini, C.C. Hagan, D. Mobbs (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States. PsyarXiv (10).3123

Yeşilkanat, C.M. (2020). Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine-learning algorithm. Chaos, Solitons & Fractals (140): 110210.

Zhang, G.Q. Public crisis management and government responsibility: Take SARS epidemic control as an example. Management World 2013, 12, 42–50.

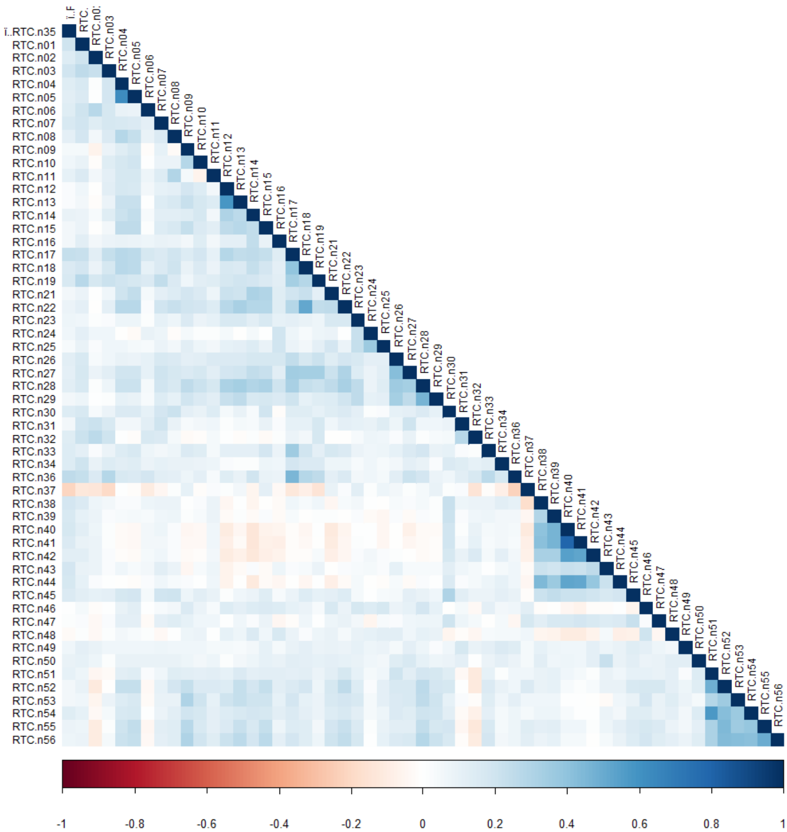
Covid Economics 65, 20 January 2021: 188-222

APPENDIX 1. Tested statements

Variable	BEHAVIOR
RTC.n1	I actively encourage others to follow the restrictions and guidelines
RTC.n2	I comply with the recommendations for physical distancing
RTC.n3	I comply with the restrictions to stay home
RTC.n4	I disinfect groceries before putting them away
RTC.n5	I disinfect mail and deliveries before opening them
RTC.n6	I wash hands for 20 seconds when necessary
RTC.n7	I would like to help people who are more vulnerable to COVID-19
RTC.n8	Since COVID-19 I eat healthier
RTC.n9	Since COVID-19 I eat unhealthier
RTC.n10	Since COVID-19 I exercise less
RTC.n11	Since COVID-19 I exercise at home more
RTC.n12	When a COVID-19 vaccine is available, I'd like to be vaccinated
	EMOTIONS
RTC.n13	I'm worried about my financial situation
RTC.n14	I'm worried about my job situation
RTC.n15	I'm worried that our country will run out of money
RTC.n16	I'm worried that there will not be enough basic necessities in the stores
RTC.n17	I am worried about my own health
RTC.n18	I am worried about the health of my children
RTC.n19	I am worried about the health of my older family members
RTC.n20	I am worried about the health of people in my country
RTC.n21	I worry that there will be an increase in break-ins and thefts
RTC.n22	I'm worried about my children's education
RTC.n23	I am anxious about not being able to meet with friends
RTC.n24	I am worried about not being able to meet with my family
RTC.n25	I worry how living in isolation will affect me
RTC.n26	Living in isolation negatively impacts my wellbeing
	OPINIONS
RTC.n27	The COVID-19 outbreak will make society more unequal
RTC.n28	Being together all the time increases family tensions
RTC.n29	COVID-19 increases domestic violence
RTC.n30	COVID-19 will increase divorce rates
RTC.n31	COVID-19 will bring countries closer
RTC.n32	I am grateful to our essential workers
RTC.n33	I am grateful to our healthcare professionals
RTC.n34	My chance of getting COVID-19 is high
RTC.n35	Slowing the spread of COVID-19 is more important than the economy
RTC.n36	Coronavirus is dangerous for my health
RTC.n37	Media exaggerate the situation with COVID-19
RTC.n38	Media provide reliable information about the pandemic
RTC.n39	[The President] is doing a good job dealing with COVID-19
RTC.n40	I am satisfied with how my government is handling this crisis
RTC.n41	The government is doing a good job dealing with COVID-19
RTC.n42	I am satisfied with how our healthcare system is handling this crisis
RTC.n43	In the case of coronavirus infection, I will get appropriate medical help
RTC.n44	The government discloses real numbers of coronavirus infections and deaths
RTC.n45	COVID-19 reveals the best in people
RTC.n46	COVID-19 reveals the worse in people
RTC.n47	I believe we will beat COVID-19 soon
RTC.n48	People will stop following the restrictions soon

Covid Economics 65, 20 January 2021: 188-222

APPENDIX 2. Correlation matrix with vaccination acceptance



Notes:

The correlation matrix shows that vaccination intent is highly negatively correlated with "Media exaggerate the situation with COVID-19" and highly positively correlated with the following statements:

- I actively encourage others to follow the restrictions;
- I comply with the restrictions to stay home;
- I'd like to help people who are more vulnerable to COVID-19;
- I am worried about my own health;
- I am worried about the health of my children;
- I am worried about the health of my older family members;
- COVID-19 will bring countries closer;
- Coronavirus is dangerous for my health;
- In case of an infection, I will get appropriate medical help.

APPENDIX 3a. Logistic regression by “acceptors”, “hesitant”, and “refusers”

VARIABLES	(1) accepting	(2) se	(3) hesitates	(4) se	(5) refuses	(6) se
Spain	0.298**	(0.151)	-0.0704	(0.160)	-0.469**	(0.214)
France	-0.503***	(0.135)	0.447***	(0.143)	0.229	(0.181)
Italy	0.360***	(0.139)	-0.146	(0.146)	-0.388*	(0.202)
Sweden	-0.628***	(0.135)	0.509***	(0.142)	0.357**	(0.181)
The UK	0.054	(0.131)	0.206	(0.140)	-0.457***	(0.180)
Female	-0.0359	(0.464)	0.352	(0.440)	-0.478	(0.526)
Male	0.292	(0.464)	0.0786	(0.441)	-0.603	(0.528)
26-35	-0.289	(0.200)	0.433**	(0.212)	-0.0139	(0.285)
36-49	-0.583***	(0.174)	0.491***	(0.189)	0.333	(0.252)
50-64	-0.662***	(0.164)	0.490***	(0.180)	0.429*	(0.237)
>64	-0.519***	(0.148)	0.502***	(0.163)	0.131	(0.217)
Primary school	0.0551	(0.214)	-0.0951	(0.222)	0.0766	(0.285)
Vocational	-0.0490	(0.125)	0.0148	(0.133)	0.0522	(0.167)
High school	-0.142	(0.125)	0.245*	(0.133)	-0.140	(0.170)
Bachelor or higher	-0.0328	(0.129)	0.146	(0.138)	-0.170	(0.176)
Kids 1	-0.268	(0.301)	0.400	(0.332)	-0.205	(0.320)
Kids 2	-0.0777	(0.306)	0.135	(0.336)	-0.128	(0.325)
Kids 3	-0.164	(0.306)	0.282	(0.337)	-0.189	(0.329)
Kids>3	-0.272	(0.338)	0.173	(0.368)	0.0965	(0.371)
Employed	-0.159	(0.178)	0.267	(0.183)	-0.109	(0.241)
Entrepreneur	-0.256	(0.214)	0.322	(0.221)	-0.0788	(0.293)
Unemployed	-0.176	(0.185)	0.250	(0.189)	-0.0756	(0.253)
Retired	0.146	(0.212)	0.0717	(0.222)	-0.386	(0.297)
<100,000 habitants	-0.0592	(0.0733)	0.0598	(0.0776)	0.00929	(0.102)
<2000 euros/month	-0.0198	(0.120)	-0.0397	(0.122)	0.0399	(0.164)
>2000 euros/month	0.184	(0.120)	-0.247**	(0.123)	-0.0144	(0.166)
Exposed to Cov	0.159	(0.150)	-0.345**	(0.150)	0.320	(0.224)
Not exposed	-0.0120	(0.139)	-0.208	(0.136)	0.420**	(0.208)
Does not know	0.156	(0.428)	-0.213	(0.428)	0.281	(0.522)
Left	0.265*	(0.144)	-0.351**	(0.145)	0.0547	(0.215)
Right	0.346**	(0.147)	-0.341**	(0.147)	-0.0667	(0.216)
Other	-0.153	(0.143)	-0.0248	(0.143)	0.304	(0.207)
Don't associate with politics	-0.202	(0.141)	0.112	(0.139)	0.168	(0.208)

Covid Economics 65, 20 January 2021: 188-222

APPENDIX 3a. Logistic regression by "acceptors", "hesitant", and "refusers"

VARIABLES	(1) accepting	(2) se	(3) hesitates	(4) se	(5) refuses	(6) se
rtcn07	0.734***	(0.231)	-0.660***	(0.233)	-0.235	(0.321)
rtcn08	0.0658	(0.144)	-0.0810	(0.151)	0.0657	(0.197)
rtcn09	0.144	(0.168)	0.148	(0.175)	-0.448*	(0.239)
rtcn10	-0.00312	(0.142)	-0.131	(0.149)	0.243	(0.202)
rtcn11	0.133	(0.140)	0.0247	(0.147)	-0.260	(0.198)
rtcn12	-0.184	(0.157)	0.0451	(0.159)	0.225	(0.220)
rtcn13	0.0608	(0.162)	-0.124	(0.170)	0.0991	(0.220)
rtcn14	0.218	(0.200)	-0.387*	(0.206)	0.117	(0.288)
rtcn15	0.197	(0.190)	-0.142	(0.198)	-0.105	(0.267)
rtcn16	-0.289	(0.232)	0.0547	(0.241)	0.512	(0.325)
rtcn17	0.316**	(0.150)	0.0381	(0.159)	-0.728***	(0.217)
rtcn18	0.0388	(0.166)	0.124	(0.172)	-0.156	(0.223)
rtcn19	0.313	(0.192)	0.00555	(0.201)	-0.336	(0.247)
rtcn21	-0.169	(0.173)	0.156	(0.177)	0.0511	(0.237)
rtcn22	-0.210	(0.163)	0.243	(0.166)	-0.0530	(0.230)
rtcn23	0.318	(0.197)	-0.278	(0.203)	-0.0635	(0.279)
rtcn24	0.0122	(0.180)	0.235	(0.184)	-0.379	(0.246)
rtcn25	0.160	(0.196)	-0.250	(0.198)	0.152	(0.275)
rtcn26	0.336**	(0.167)	-0.162	(0.173)	-0.327	(0.227)
rtcn27	0.0884	(0.177)	-0.0427	(0.184)	-0.0739	(0.245)
rtcn28	0.409**	(0.206)	-0.392*	(0.211)	-0.164	(0.305)
rtcn29	-0.0613	(0.186)	-0.276	(0.190)	0.455*	(0.255)
rtcn30	0.268	(0.182)	0.0919	(0.186)	-0.644**	(0.276)
rtcn31	0.245	(0.233)	-0.0511	(0.241)	-0.282	(0.312)
rtcn32	0.652***	(0.242)	-0.00140	(0.253)	-0.794***	(0.298)
rtcn33	0.284	(0.195)	0.123	(0.197)	-0.726**	(0.306)
rtcn34	1.109***	(0.270)	-0.748***	(0.280)	-0.681*	(0.374)
rtcn36	0.839***	(0.167)	-0.290*	(0.175)	-0.809***	(0.211)
rtcn37	-0.830***	(0.158)	0.165	(0.167)	1.122***	(0.215)
rtcn38	0.413**	(0.197)	0.0572	(0.205)	-0.781***	(0.274)
rtcn39	0.199	(0.184)	-0.0672	(0.190)	-0.120	(0.257)
rtcn40	0.117	(0.235)	0.0634	(0.243)	-0.342	(0.321)
rtcn41	-0.239	(0.228)	0.170	(0.232)	0.207	(0.321)
rtcn42	-0.0101	(0.196)	0.251	(0.210)	-0.347	(0.271)
rtcn43	1.187***	(0.253)	-0.707***	(0.247)	-0.976***	(0.356)
rtcn44	0.494**	(0.239)	0.156	(0.249)	-1.091***	(0.334)
rtcn45	0.171	(0.184)	-0.0298	(0.192)	-0.207	(0.283)
rtcn46	0.0645	(0.173)	-0.224	(0.177)	0.230	(0.258)
rtcn47	-0.152	(0.164)	-0.0631	(0.168)	0.235	(0.227)
rtcn48	-0.218	(0.210)	-0.0653	(0.210)	0.500*	(0.297)
rtcn49	0.299	(0.252)	-0.225	(0.259)	-0.158	(0.342)
Constant	-3.265***	(0.831)	-0.395	(0.800)	2.254**	(1.014)
Observations	6,155		6,155		6,155	

Notes: See Table 1 and Appendix 1 for default and label of statements.
Prob > F= 0.000, root mean square error = 0.15, pseudo R^2 = 0.23.

APPENDIX 3b. Computed marginal propensity to vaccinate, multi-logit estimates (non adjusted by response time)

Statements	Probability
In case of a coronavirus infection, I will get appropriate medical help	10.1%
Coronavirus is dangerous for my health	9.8%
I am worried about my health	6.7%
The government discloses real numbers of coronavirus infections and deaths	6.5%
[PRESIDENT] is doing a good job dealing with COVID-19	5.8%
I disinfect mail and deliveries before opening them	5.4%
My chance of getting COVID-19 is high	5.3%
I am satisfied with how my government is handling this crisis	5.0%
I comply with the restrictions to stay home	4.8%
I actively encourage others to follow the restrictions and guidelines	4.7%
I would like to help people who are more vulnerable to COVID	4.4%
Slowing the spread of COVID-19 is more important than the economy	3.9%
I am anxious about not being able to meet with friends	3.9%
I am grateful to our healthcare professionals	3.6%
I comply with the recommendations for physical distancing	3.6%
COVID-19 will bring countries closer	3.5%
I wash hands for 20 seconds when necessary	3.5%
I am worried about the health of people in my country	3.3%
Being together all the time increases family tensions	3.2%
I worry how living in isolation will affect me	3.2%
I am worried about my financial situation	-3.8%
I am grateful to our essential workers	-4.6%
Media exaggerate the situation with COVID-19	-13.1%

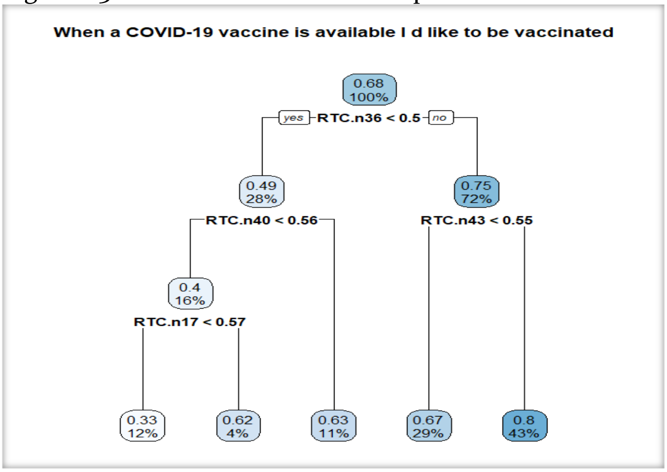
Notes: Only statistically significant parameters are shown ($\alpha<5\%$). Multilogit on three categories (acceptors, hesitant, and refuses). Country dummy included, as well as socio-demographics.

Covid Economics 65, 20 January 2021: 188-222

Covid Economics 65, 20 January 2021: 188-222

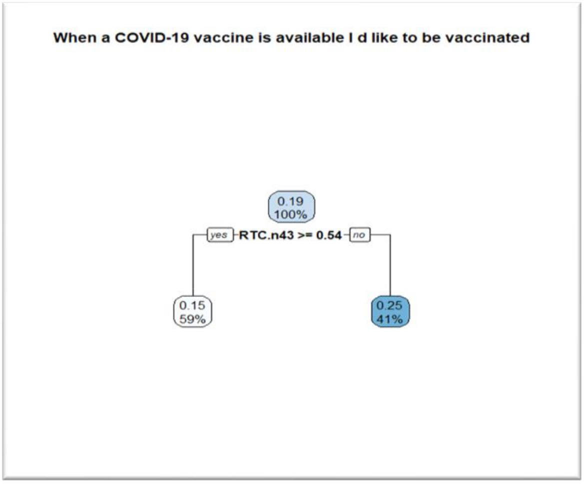
APPENDIX 4. Classification trees by "acceptors", "hesitant", and "refusers" (non-adjusted by response time)

Figure A3.1. Classification tree -Acceptor



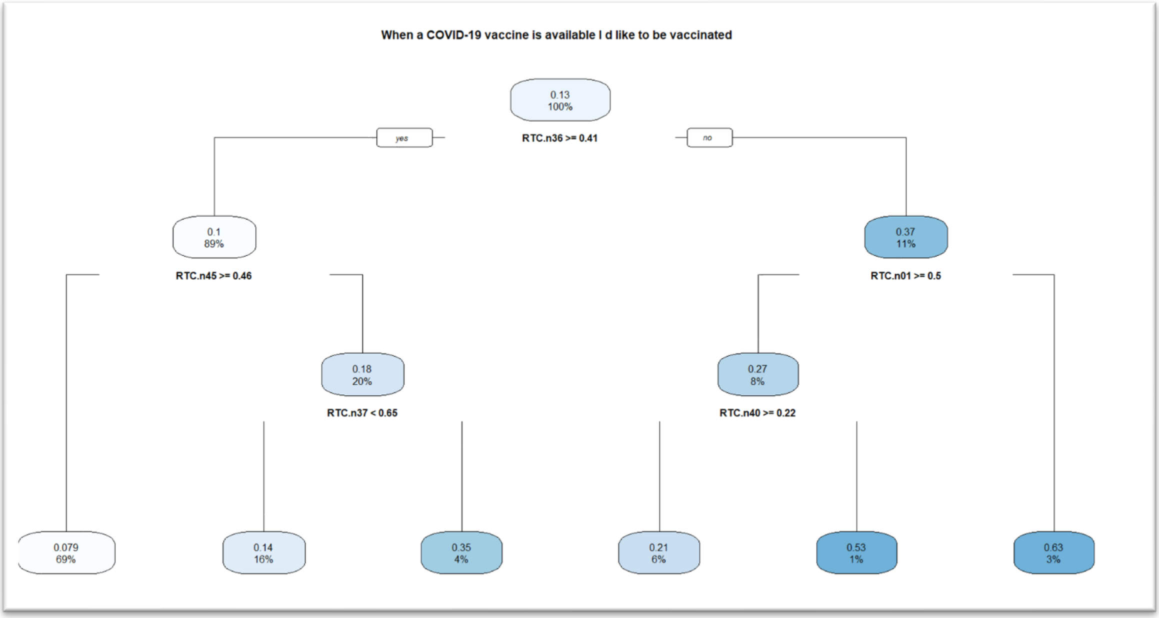
Notes:
RTC.n36: Coronavirus is dangerous for my health
RTC.n17: I am worried about my own health
RTC.n 40: I am satisfied with how my government is handling this crisis
RTC.n43: In case of a coronavirus infection I will get appropriate medical help

Figure A3.2. Classification tree - Hesitant



Note:
RTC.n43: In case of a coronavirus infection I will get appropriate medical help.

Figure A3.3. Classification tree - Refusers



Notes:
RTC.n1: I actively encourage others to follow the restrictions and guidelines;
RTC.n2: I'm worried about my children's education;
RTC.n37: Media exaggerate the situation with Covid-19;
RTC.n0: The government is doing a good job dealing with Covid-19;
RTC.n46: Covid-19 reveals the worse in people.

Covid Economics 65, 20 January 2021: 188-222